

# Décomposition M-Vine de processus vecteur autorégressif d'ordre 1

par

Étienne Bégin

mémoire présenté au Département de mathématiques  
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, mai 2019

Le 14 mai

*le jury a accepté le mémoire de Monsieur Étienne Bégin  
dans sa version finale.*

Membres du jury

Professeur Taoufik Bouezmarni  
Directeur de recherche  
Département de Mathématiques

Professeur Carole Beaulieu  
Codirectrice de recherche  
Département de Biologie

Professeur Pierre Dutilleul  
Codirecteur de recherche  
McGill University, Department of Mathematics and Statistics

Professeur Félix Camirand Lemyre  
Membre interne  
Département de Mathématiques

Professeur Éric Marchand  
Président-rapporteur  
Département de Mathématiques

Pour ma conjointe Laurence,  
qui m'a soutenu du début à la fin,  
ainsi que mes enfants  
Lino et Automne.

# SOMMAIRE

La modélisation de structures de dépendance est un concept récurrent en statistique. Ce mémoire veut étudier une nouvelle façon d’aborder et de modéliser la dépendance pour des séries chronologiques multivariées.

Le chapitre 1 aborde les séries chronologiques : des caractéristiques propres aux séries temporelles et quelques familles de séries chronologiques y sont mentionnées et décrites.

Le chapitre 2 se penche sur l’étude de la structure de dépendance, et plus particulièrement sur les copules. On y définit le concept de copules pour modéliser des structures de dépendance et ses propriétés, quelques classes et familles de copules et la décomposition de type Vine pour les variables multivariées.

Le chapitre 3 étudie la relation entre la modélisation de la dépendance du modèle vecteur autorégressif d’ordre 1 ( $\text{VAR}(1)$ ) et celle issue de la décomposition M-Vine. On y décrit les manipulations pour passer d’une représentation  $\text{VAR}(1)$  à une représentation M-Vine.

Le chapitre 4 montre deux applications du passage du  $\text{VAR}(1)$  au M-Vine. La première est un exemple théorique issu de Beare et Seo (2015) [BS15]. La deuxième est une application empirique qui s’intéresse à la structure de dépendance de prix du blé dans trois pays.

Mots-clés : Décomposition M-Vine, Processus vecteur autorégressif, Copule Gaussienne Multivariée, Séries chronologiques, Copules, Structure de dépendance

# REMERCIEMENTS

Je tiens d'abord à remercier mon directeur de maîtrise, M. Taoufik Bouezmarni, pour le soutien qu'il m'a apporté durant ma maîtrise et pour les nombreuses questions auxquelles il m'a répondu. Je veux aussi souligner l'apport de mon directeur de recherche adjoint, M. Pierre Dutilleul, qui m'a beaucoup aidé à structurer et expliciter mes idées pour bien les présenter. Je veux remercier Mme Carole Beaulieu pour les échanges constructifs qu'on a pu avoir et son point de vue "moins" mathématiques qui m'a fait réfléchir et progresser. Ce mémoire a été rendu possible grâce au soutien financier principalement de AgroPhytoSciences du Centre SÈVE, mais aussi de l'ISM, de la famille Audet et de la fondation FORCE.

Étienne Bégin

Sherbrooke, mai 2019

# TABLE DES MATIÈRES

<b>SOMMAIRE</b>	<b>iii</b>
<b>REMERCIEMENTS</b>	<b>iv</b>
<b>TABLE DES MATIÈRES</b>	<b>v</b>
<b>LISTE DES TABLEAUX</b>	<b>viii</b>
<b>LISTE DES FIGURES</b>	<b>ix</b>
<b>INTRODUCTION</b>	<b>1</b>
<b>CHAPITRE 1 — Séries chronologiques</b>	<b>5</b>
1.1 Modèles classiques . . . . .	8
1.1.1 MA . . . . .	9
1.1.2 AR . . . . .	11
1.1.3 ARMA . . . . .	13

1.1.4	VAR . . . . .	15
1.1.5	(G)ARCH . . . . .	17
<b>CHAPITRE 2 — Copules</b>		<b>21</b>
2.1	Définition et propriétés . . . . .	23
2.2	Classes et familles de copules . . . . .	25
2.2.1	Classe elliptique . . . . .	25
2.2.2	Classe archimédienne . . . . .	26
2.2.3	Copule d'indépendance . . . . .	26
2.2.4	Copule gaussienne . . . . .	27
2.2.5	Copule de Student . . . . .	29
2.2.6	Copule de Clayton . . . . .	30
2.2.7	Copule de Frank . . . . .	31
2.2.8	Copule de Gumbel . . . . .	32
2.2.9	Copule d'Ali-Mikhail-Haq (AMH) . . . . .	33
2.3	Méthode d'estimation des copules . . . . .	34
2.3.1	Estimation paramétrique . . . . .	35
2.3.2	Estimation semi-paramétrique . . . . .	36
2.3.3	Estimation non-paramétrique . . . . .	36
2.4	Décomposition de type Vine . . . . .	37
2.4.1	C-Vine et D-Vine . . . . .	38

2.4.2	M-Vine . . . . .	44
<b>CHAPITRE 3 — Lien entre VAR(1) et M-Vine</b>		<b>47</b>
3.1	Représentation de la structure de dépendance d'un VAR(1) avec une copule gaussienne multivariée . . . . .	49
3.2	Décomposition M-Vine d'une copule gaussienne multivariée . . . . .	53
3.3	Recomposition d'une copule gaussienne multivariée à partir de sa décomposition M-Vine . . . . .	58
3.4	Retour au modèle classique VAR(1) à partir d'une copule gaussienne multivariée . . . . .	64
<b>CHAPITRE 4 — Applications</b>		<b>66</b>
4.1	Application théorique . . . . .	67
4.2	Application empirique : prix du blé . . . . .	69
<b>CONCLUSION</b>		<b>76</b>
<b>BIBLIOGRAPHIE</b>		<b>77</b>



# LISTE DES TABLEAUX

4.1	Paramètres estimés pour une représentation M-V des prix du blé utilisant des copules gaussiennes . . . . .	72
4.2	Résultats pour une représentation M-V des prix du blé en utilisant des copules à partir d'un ensemble de familles préselectionnées . . . . .	75

# LISTE DES FIGURES

1.1	Graphique d'une réalisation partielle de 500 observations d'un processus MA(3) de paramètres $\tau_1 = 1$ , $\tau_2 = -0,5$ , $\tau_3 = 0.4$ , $\sigma_\epsilon = \sqrt{0,25}$ et $\mu = 0$ et de sa fonction d'autocorrélation . . . . .	10
1.2	Graphique d'une réalisation partielle de 500 observations d'un processus AR(3) de paramètres $\theta_1 = 1$ , $\theta_2 = -0,5$ , $\theta_3 = 0.4$ , $\sigma_\epsilon = \sqrt{0,25}$ et $v = 0$ et de sa fonction d'autocorrélation partielle . . . . .	12
1.3	Graphique d'une réalisation partielle de 500 observations d'un processus ARMA(3,3) de paramètres $\tau_1 = \theta_1 = 1$ , $\tau_2 = \theta_2 = -0,5$ , $\tau_3 = \theta_3 = 0.4$ , $\sigma_\epsilon = \sqrt{0,25}$ et $\mu = v = 0$ . . . . .	14
1.4	Graphiques des fonctions d'autocorrélation et d'autocorrélation partielle d'une réalisation partielle de 500 observations d'un processus ARMA(3,3) de paramètres $\tau_1 = \theta_1 = 1$ , $\tau_2 = \theta_2 = -0,5$ , $\tau_3 = \theta_3 = 0.4$ , $\sigma_\epsilon = \sqrt{0,25}$ et $\mu = v = 0$ . . . . .	15
1.5	Graphique d'une réalisation partielle de 500 observations d'un processus ARCH(2) de paramètres $\alpha_0 = 0$ , $\alpha_1 = 0.2$ et $\alpha_2 = 0.4$ . . . . .	19
1.6	Graphique d'une réalisation partielle de 500 observations d'un processus GARCH(1,1) de paramètres $\alpha_0 = 0$ , $\alpha_1 = 0.2$ et $\beta_1 = 0.7$ . . . . .	20

2.1	Graphique de densité et nuage de points d'une copule d'indépendance bivariée $C(u, v)$ . . . . .	27
2.2	Graphiques de densité et nuages de points de copules gaussiennes bivariées $C(u, v; \rho)$ avec $\rho = -0,75$ , $\rho = 0$ et $\rho = 0,75$ . . . . .	28
2.3	Graphiques de densité et nuages de points de copules de Student bivariées $C(u, v; \rho, \kappa = 5)$ avec $\rho = -0,75$ , $\rho = 0$ et $\rho = 0,75$ . . . . .	30
2.4	Graphique de densité et nuage de points d'une copule de Clayton bivariée $C_\theta(u, v)$ avec $\theta = 3$ . . . . .	31
2.5	Graphique de densité et nuage de points d'une copule de Frank bivariée $C_\theta(u, v)$ avec $\theta = 7$ . . . . .	32
2.6	Graphique de densité et nuage de points d'une copule de Gumbel bivariée $C_\theta(u, v)$ avec $\theta = 3$ . . . . .	33
2.7	Graphiques de densité et nuages de points de copules d'Ali-Mikhail-Haq bivariées $C_\theta(u, v)$ avec $\theta = 0.10$ et $\theta = 0.25$ . . . . .	34
2.8	Arbres de la décomposition D-Vine d'une copule $k$ -variée avec $k = 5$ . . .	41
2.9	Arbres de la décomposition C-Vine d'une copule $k$ -variée avec $k = 5$ . . .	43
2.10	Les arbres $T_1$ , $T_2$ , $T_3$ , $T_7$ et $T_8$ de la décomposition M-Vine d'une copule associée à une série chronologique markovienne d'ordre 3 tri-variée . . . .	46
3.1	Schéma des transitions de représentation avec leur section associée . . . .	48
4.1	Séries chronologiques des prix du blé entre novembre 2010 et avril 2014 en Australie, au Kazakhstan et au Royaume-Uni . . . . .	69

# INTRODUCTION

Statisticiens et mathématiciens veulent toujours contruire des modèles pour représenter des phénomènes observables. L'un de ces phénomènes est la variation de mesures dans le temps. Les observations de ces mesures peuvent être réunies sous la forme d'une série, qu'on appelle série chronologique. Un observateur peut alors s'intéresser à une réalisation partielle à temps discret de ce processus stochastique temporel. C'est avec un souci mathématique de modélisation de ces séries chronologiques que Yule (1927) [UY27] et Walker (1931) [Wal31] proposent le premier modèle retenu pour représenter de telles séries. Ils nomment cette représentation, le modèle autorégressif (AR). Un AR d'ordre  $p$  consiste à expliquer une valeur à un temps  $t$  par les  $p$  valeurs précédentes desquelles on ajoute une innovation. Ce modèle fut ensuite généralisé au cas multivarié et est noté VAR (vecteur autorégressif). AR et VAR sont dits modèles classiques pour l'analyse de séries chronologiques.

Bien entendu, un seul type de modèle n'est pas suffisant pour représenter toutes les séries chronologiques. Whittle (1951) [Whi51] aborde la représentation moyenne mobile ("moving average" ou MA) pour les séries chronologiques qui consiste à expliquer la valeur prise au temps  $t$  par les innovations aux temps  $t, (t - 1), \dots, (t - q)$  ( $q$  est alors l'ordre de la moyenne mobile). La combinaison des modèles AR et MA forme la représentation ARMA. Cette dernière correspond au modèle autorégressif moyenne mobile d'ordre  $p$  et

$q$ , dans lequel la valeur d'une série au temps  $t$  est expliquée par les valeurs prises aux temps  $(t - i)$  et par les innovations aux temps  $(t - j)$  avec  $i = 1, \dots, p$  et  $j = 0, \dots, q$ .

Bien que les modèles AR, MA et ARMA ajustent très bien au comportement de plusieurs phénomènes aléatoires variant dans le temps, ils ne peuvent pas modéliser la présence de volatilité dans les valeurs observées. Engle (1982) [Eng82] propose alors le modèle ARCH (autorégressif conditionnellement hétéroscédastique) pour représenter des séries chronologiques dont la variance n'est pas stable dans le temps. Par la suite, Bollerslev (1986) [Bol86] a généralisé le tout pour obtenir le modèle GARCH (autorégressif conditionnellement hétéroscédastique généralisé).

Beaucoup d'autres modèles ont été introduits depuis, mais AR (VAR), MA, ARMA, ARCH et GARCH sont, sans contredit, les plus utilisés pour représenter les séries chronologiques. Peu importe le modèle utilisé pour représenter un phénomène aléatoire variant dans le temps, la structure de dépendance entre les observations d'une série en des temps différents, et dans la cas multivarié, entre les séries, se résume souvent à une seule valeur, comme la corrélation, la covariance ou à une seule matrice (matrice de variance-covariance). Cette seule valeur (ou matrice de valeurs) pour mesurer la dépendance peut parfois s'avérer insuffisante pour capturer certaines subtilités dans un modèle. Cependant, un autre outil permet d'étudier la structure de dépendance de façon plus flexible : les copules.

Les copules permettent d'aborder la dépendance entre des variables aléatoires comme une loi de probabilité. Ainsi, pour quantifier la dépendance, on ne s'intéresse plus à une valeur de corrélation, mais à une distribution de probabilité. En 1959, c'est Sklar [Skl59] qui introduit le concept de copules. Schweizer et Sklar (1983) [SS83] sont parmi les premiers à utiliser le concept de copule après la formulation du concept par Sklar [Skl59]. Les copules deviennent plus populaires vers la fin des années 90 et au début des années 2000. Nelsen (1999) [Nel99], Joe (1997) [Joe97], Drouet, Mari et Kotz (2001) [MK01] et Genest et

Rivest (1993) [GR93] pavent la voie aux statisticiens et aux chercheurs pour l'étude et l'utilisation des copules. Dans le monde de la finance, Embrechts, McNeil et Straumann (2002) [EMS02], Frees et Valdez (1998) [FV98] et Patton (2006) [Pat06] ont apportés une grande contribution pour l'utilisation des copules dans un cadre économique et financier.

L'un des principaux défauts des copules provient de son utilisation pour les vecteurs aléatoires de grandes dimensions. En effet, ajuster la structure de dépendance d'une variable aléatoire multidimensionnelle à une copule multivariée peut s'avérer fastidieux. De plus, en se restraignant aux copules paramétriques, l'ajustement est souvent très mauvais. C'est dans l'optique de résoudre ce problème qu'on en vient à décrire la décomposition de type Vine de copules multivariées. L'objectif de la décomposition de type Vine est de représenter la densité d'une copule multivariée sous forme de produit de densités de copules bivariées. En 1996, Joe [Joe96] est le premier à décrire la décomposition de type Vine en introduisant la décomposition D-Vine. Bedford et Cooke (2002) [BC02] et Kurowicka et Cooke (2006) [KC06] introduisant quant à eux la décomposition R-Vine et C-Vine respectivement. L'intérêt de ces décompositions est de pouvoir estimer une copule multivariée à l'aide de plusieurs copules bivariées. Aas et associés (2009) [ACFB09] aborde la reconstruction d'une copule multivariée à partir des copules bivariées obtenues à l'aide d'une décomposition Vine.

En bref, les décompositions de copules s'avèrent un outil flexible donnant plus d'informations sur la structure de dépendance de variables aléatoires que peut le faire une valeur de corrélation ou une matrice de variance-covariance. L'utilisation de ces décompositions de copules pour les séries chronologiques a amené, en 2015, Beare et Seo [BS15] à introduire la décomposition M-Vine. Le M-Vine permet la construction d'une représentation de séries chronologiques plus générale que peut le faire un modèle VAR(1). Dans l'article de Beare et Seo (2015) [BS15], ceux-ci affirment que le modèle VAR(1) est un cas particulier du M-Vine sans toutefois montrer comment représenter un modèle VAR(1) quelconque

comme une décomposition M-Vine. Il s'agit ici de la raison d'être de ce travail : on cherche à étudier les liens entre un VAR(1) et une décomposition M-Vine.

Dans le but d'étudier ces liens et, ultimement, de pouvoir écrire un VAR(1) comme une décomposition M-Vine, au chapitre 1, une revue de certains modèles classiques pour les séries chronologiques est faite. Ensuite, au chapitre 2, les notions et concepts relatifs aux copules et à quelques décompositions de type Vine sont abordés. Le chapitre 3 s'attaque par la suite au passage d'une représentation VAR(1) d'une série chronologique multivariée à une représentation d'une décomposition M-Vine. L'étude de ce passage permet de faire des liens entre les deux modèles. Un manuscrit qui résume les contributions de ce troisième chapitre est soumis pour publication à la revue *Statistics and Probability Letters*. Ce manuscrit est reproduit en Annexe. Finalement, au chapitre 4, des applications théorique et empirique sont présentées.

# CHAPITRE 1

## Séries chronologiques

Un sujet bien étudié en statistiques est celui des séries chronologiques, ou suite de variables aléatoires variant dans le temps. Plusieurs ouvrages de références sont disponibles aux chercheurs s'intéressant au sujet. Pour ne nommer que quelques livres : Priestley (1989) [Pri89], Hamilton (1994) [Ham94], Enders (1995) [End95] et Brockwell et Davis (1991) [BD91]. Ce chapitre est principalement basé sur les résultats exposés dans ces derniers ouvrages de référence. Avant d'étudier en détail les différents modèles de séries chronologiques, certaines définitions et propriétés fréquemment utilisées se doivent d'être introduites.

Tout d'abord, on rappelle qu'une série chronologique est un processus stochastique  $(X_t)_{t \in \mathbb{Z}}$ , où  $t$  représente le temps. De façon générale, un processus  $(X_t)_{t \in \mathbb{Z}}$  dépend, d'une certaine façon, des observations précédentes  $X_s$ ,  $s < t$ . Une notion à connaître lors de l'étude des séries chronologiques est le concept de stationnarité. La stationnarité est une caractéristique d'une série chronologique dont la dépendance ne dépend pas du temps. Formellement, on différencie deux types de stationnarité : faible et forte.

Une série  $(X_t)_{t \in \mathbb{Z}}$  est dite :



1. faiblement stationnaire si  $E[X_t]$  et  $\gamma(l)$ ,  $\forall l \in \mathbb{Z}$ , sont finies et ne dépendent pas du temps  $t$  ;
2. fortement stationnaire si  $\forall t \in \mathbb{Z}$ , la loi de  $(X_1, X_2, \dots, X_t)$  est la même que la loi de  $(X_{1+l}, X_{2+l}, \dots, X_{t+l})$   $\forall l \in \mathbb{Z}$ .

Aussi, on note que  $(X_t)_{t \in \mathbb{Z}}$  est stable si elle ne comporte pas de tendance saisonnière, ni de tendance à la hausse ou à la baisse. Notons qu'à chaque fois qu'on discutera de la stationnarité d'un processus dans ce mémoire, on fera référence à la stationnarité faible.

Ensuite, pour mieux étudier les liens d'interdépendance entre les termes de cette série, on utilise les fonctions d'autocovariance, d'autocorrélation et d'autocorrélation partielle qui, comme leur nom l'indique, mesurent la variance et la corrélation entre les éléments d'une même série. On note que, pour une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  faiblement stationnaire, la fonction d'autocovariance de délai  $l$ ,  $\forall l \in \mathbb{Z}$ , est définie par :

$$\gamma(l) = cov(X_t, X_{t-l}). \quad (1.1)$$

On constate qu'on a  $\gamma(0) = var(X_t)$ . La fonction d'autocorrélation de délai  $l$ , quant à elle, est définie à partir de la fonction d'autocovariance comme suit :

$$\rho(l) = cor(X_t, X_{t-l}) = \frac{\gamma(l)}{\gamma(0)}. \quad (1.2)$$

On s'intéresse à ces dernières fonctions puisqu'elles fournissent de l'information sur la structure de dépendance d'une série ; structure avec laquelle on veut, dans le chapitre 3, présenter une relation avec les copules et leur décomposition de type Vine (concepts abordés dans le chapitre 2).

Le concept d'autocorrélation partielle s'intéresse aussi à la corrélation entre des termes  $X_t$  et  $X_{t-l}$  d'une série  $(X_t)_{t \in \mathbb{Z}}$ . Cependant, elle ne prend pas en considération l'influence des variables intermédiaires entre  $X_t$  et  $X_{t-l}$ , c'est-à-dire les variables  $X_{t-l+h}$  avec  $h = 1, 2, \dots, (l-1)$ . Avec  $det(\cdot)$  l'opération retournant le déterminant d'une matrice, il est

possible de montrer que cette autocorrélation partielle de délai  $l \in \mathbb{N}$  d'un processus  $(X_t)_{t \in \mathbb{Z}}$  est donnée par

$$\rho^*(l) = \frac{\det(P^*(l))}{\det(P(l))}, \quad (1.3)$$

où

$$P^*(l) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(l-2) & \rho(1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(l-3) & \rho(2) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho(l-1) & \rho(l-2) & \rho(l-3) & \cdots & \rho(1) & \rho(l) \end{pmatrix}$$

et

$$P(l) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(l-2) & \rho(l-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(l-3) & \rho(l-2) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho(l-1) & \rho(l-2) & \rho(l-3) & \cdots & \rho(1) & 1 \end{pmatrix}.$$

Cette dernière fonction est forte utile pour déterminer l'ordre d'un processus autorégressif sur lequel nous reviendrons plus loin dans le présent chapitre.

Un processus utile en travaillant avec les séries chronologiques est celui du bruit blanc. Le bruit blanc  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un processus faiblement stationnaire ayant certaines propriétés supplémentaires. En plus des propriétés de stationnarité ( $E[\epsilon_t]$  et  $\gamma(l)$ ,  $\forall l \in \mathbb{Z}$ , ne dépendent pas du temps  $t$ ), on a aussi que  $E[\epsilon_t] = 0$ ,  $E[\epsilon_t^2] < \infty$  et  $\rho(l) = 0 \forall l \geq 1$ . Ainsi, un bruit blanc au temps  $t$  n'a aucune incidence sur ce même bruit blanc au temps  $t + l \forall l \in \mathbb{Z}$ . On dit qu'un bruit blanc est gaussien lorsque  $\dots, \epsilon_{-1}, \epsilon_0, \epsilon_1, \dots$  sont i.i.d.  $N(0, \sigma_\epsilon^2)$ .

Connaissant la définition d'un bruit blanc, on peut décrire l'inversibilité et la causalité d'une série chronologique. Soit  $\epsilon_t$  un bruit blanc, alors un processus  $(X_t)_{t \in \mathbb{Z}}$  est dit :

1. inversible s'il est possible d'écrire un bruit blanc comme une combinaison linéaire des itérations du processus. C'est-à-dire,  $(X_t)_{t \in \mathbb{Z}}$  est inversible si et seulement s'il

existe  $\alpha_k \in \mathbb{R}$  avec  $k \in \mathbb{N}$  et  $\sum_{k=0}^{\infty} |\alpha_k| < \infty$  tels que

$$\epsilon_t = \sum_{k=0}^{\infty} \alpha_k X_{t-k} \quad (1.4)$$

2. causal s'il est possible d'écrire le processus comme une combinaison linéaire d'itérations d'un bruit blanc. C'est-à-dire,  $(X_t)_{t \in \mathbb{Z}}$  est causal si et seulement s'il existe  $\beta_k \in \mathbb{R}$  avec  $k \in \mathbb{N}$  et  $\sum_{k=0}^{\infty} |\beta_k| < \infty$  tels que

$$X_t = \sum_{k=0}^{\infty} \beta_k \epsilon_{t-k}. \quad (1.5)$$

Enfin, on définit aussi un opérateur délai qui permet d'écrire plus simplement certains processus à partir des réalisations passées. Pour un processus stochastique  $(X_t)_{t \in \mathbb{Z}}$ , on note cet opérateur  $B$  qui est défini par  $BX_t = X_{t-1}$ .

Prendre note qu'on s'intéresse seulement aux processus qui sont stables et faiblement stationnaires et aux bruits blancs gaussiens dans ce mémoire.

Les sections et sous-sections qui suivent s'intéressent à différents modèles définis pour étudier les séries chronologiques. Les modèles classiques ARMA (autorégressif moyenne mobile), VAR (vecteur autorégressif) et (G)ARCH (autorégressif conditionnellement hétéroscédastique (généralisé)) sont ceux qui sont abordés en détail dans ce mémoire.

## 1.1 Modèles classiques

Les modèles classiques abordés dans cette section sont ARMA, VAR et (G)ARCH. Le modèle ARMA s'avère un mélange entre les modèles AR (autorégressif) et MA (moyenne mobile) et il sera alors justifié d'aborder ces deux sous-modèles avant d'étudier le modèle mixte ARMA. Le modèle VAR est une extension multidimensionnelle du modèle AR. Quant au modèle (G)ARCH, il est employé couramment dans la modélisation de

séries temporelles, qui comportent des volatilités variables, c'est-à-dire que la variance conditionnelle au temps  $t$  est variable.

### 1.1.1 MA

On dit que le processus  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation de processus moyenne mobile d'ordre  $q$  (MA( $q$ )) si et seulement si

$$X_t = \mu + \Gamma(B) \epsilon_t \quad (1.6)$$

où

$$\Gamma(B) = \sum_{j=0}^q \tau_j B^j, \quad (1.7)$$

$(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc gaussien et  $\tau_0, \tau_1, \dots, \tau_q \in \mathbb{R}$  dont on a, par convention,  $\tau_q \neq 0$  et  $\tau_0 = 1$ .

Ainsi, pour ce processus MA( $q$ ), on a :

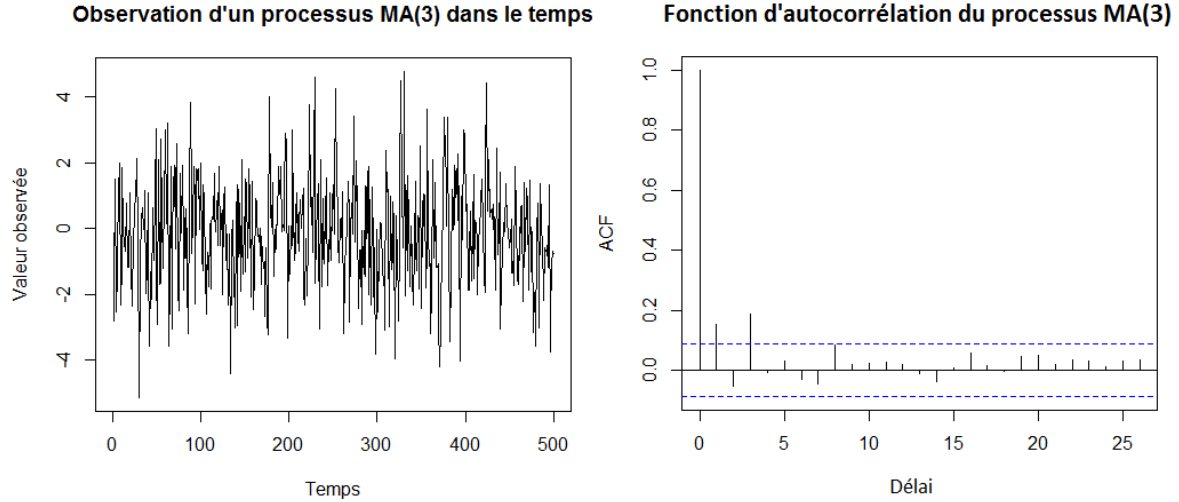
$$\begin{aligned} E[X_t] &= E[\mu + \Gamma(B) \epsilon_t] \\ &= \mu + \Gamma(B) E[\epsilon_t] = \mu \\ \gamma(l) &= \text{cov}(X_t, X_{t-l}) \\ &= \text{cov}(\mu + \Gamma(B) \epsilon_t, \mu + \Gamma(B) \epsilon_{t-l}) \\ &= \begin{cases} \sum_{i=0}^q \tau_i^2, & \text{si } l = 0 \\ \tau_l + \tau_1 \tau_{l+1} + \dots + \tau_{q-l} \tau_q, & \text{si } 0 < l \leq q \\ 0, & \text{si } l > q \end{cases} \end{aligned}$$

et donc,

$$\rho(l) = \begin{cases} 1, & \text{si } l = 0 \\ \frac{\tau_l + \tau_1 \tau_{l+1} + \dots + \tau_{q-l} \tau_q}{\sum_{i=0}^q \tau_i^2} & \text{si } 0 < l \leq q \\ 0 & \text{si } l > q \end{cases}.$$

On constate que ce processus est stationnaire peu importe les valeurs de  $\tau_i$ ,  $i = 0, 1, \dots, q$  puisque l'espérance et la fonction d'autocovariance ne dépendent pas du temps  $t$ . De plus,

Figure 1.1 – Graphique d’une réalisation partielle de 500 observations d’un processus MA(3) de paramètres  $\tau_1 = 1$ ,  $\tau_2 = -0,5$ ,  $\tau_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $\mu = 0$  et de sa fonction d’autocorrélation



on note que  $(X_t)_{t \in \mathbb{Z}}$  est toujours causal et que le processus est inversible si les racines de  $\Gamma(x)$  sont de modules strictement supérieurs à 1 [BD91].

Dans le cadre d’études empiriques, la fonction d’autocorrélation  $\rho(\cdot)$  est un outil pour estimer l’ordre  $q$  du modèle MA. L’ordre  $q$  correspondra au plus grand délai tel que le coefficient d’autocorrélation est significativement différent de zéro ; voir l’exemple à la figure 1.1. À partir d’un exemple d’un processus MA(3) de paramètres  $\tau_1 = 1$ ,  $\tau_2 = -0,5$ ,  $\tau_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $\mu = 0$ , où  $\sigma_\epsilon$  est l’écart-type du bruit gaussien  $\epsilon_t$ , illustrée à la figure 1.1 (figure de gauche), on obtient la fonction d’autocorrélation (acf) illustrée à la même figure 1.1 (graphique de droite). On constate que les autocorrélations pourront être considéré statistiquement nulles à partir du délai  $l = 4$  d’où on considère un ordre de  $q = 3$  pour le processus MA.

Les critères d’information d’Akaike (1974) [Aka74] (AIC) et de Schwarz (1978) [Sch78] (SBC) sont aussi d’excellents outils pour estimer l’ordre d’un modèle MA. En plus d’es-

timer l'ordre, ces critères aident à déterminer le type de processus. On réfère le lecteur à ces sources pour de plus amples détails sur les critères d'information.

### 1.1.2 AR

On dit que le processus  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation d'un processus autorégressif d'ordre  $p$  (AR( $p$ )) si et seulement si

$$\Theta(B) X_t = v + \epsilon_t \quad (1.8)$$

où  $v \in \mathbb{R}$ ,

$$\Theta(B) = \sum_{j=0}^p \theta_j B^j, \quad (1.9)$$

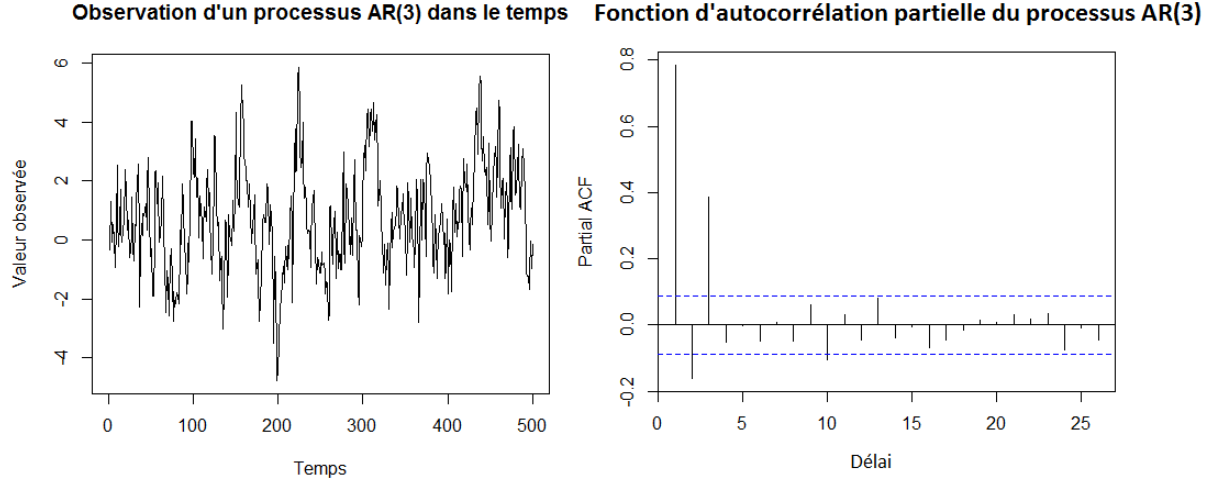
$(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc gaussien et  $\theta_0, \theta_1, \dots, \theta_p \in \mathbb{R}$  dont on a, par convention,  $\theta_p \neq 0$  et  $\theta_0 = 1$ .

En supposant  $E[X_t] = E[X_{t-l}]$  pour  $l = 0, 1, \dots, p$ , on obtient que :

$$\begin{aligned} E[X_t] &= v / \sum_{j=0}^p \theta_j \\ \text{et} \\ \rho(l) &= \begin{cases} 1, & \text{si } l = 0 \\ \sum_{j=1}^{l-p+1} \theta_j \rho(l-j), & \text{si } l \in \mathbb{Z}^* \end{cases} \end{aligned}$$

Un processus AR( $p$ ) quelconque est toujours inversible car il est toujours possible d'écrire  $X_t$  comme une combinaison linéaire des itérations d'un bruit blanc. Autrement dit, un processus autorégressif peut s'écrire comme un processus moyenne mobile d'ordre infini.

Figure 1.2 – Graphique d’une réalisation partielle de 500 observations d’un processus AR(3) de paramètres  $\theta_1 = 1$ ,  $\theta_2 = -0,5$ ,  $\theta_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $v = 0$  et de sa fonction d’autocorrélation partielle



Prenons l’exemple d’un AR(1) de la forme  $(1 - aB) X_t = v + \epsilon_t$ , alors

$$\begin{aligned}
 X_t &= v + aX_{t-1} + \epsilon_t \\
 &= v + a(v + aX_{t-2} + \epsilon_{t-1}) + \epsilon_t \\
 &= \dots \\
 &= \sum_{i=0}^{\infty} a^i v + \sum_{i=0}^{\infty} a^i \epsilon_{t-i},
 \end{aligned}$$

d’où on déduit que ce processus AR(1) peut être réécrit comme un MA( $\infty$ ). Un travail semblable peut être effectué pour un processus AR( $p$ ) général.

Cependant, un processus AR( $p$ ) n’est pas toujours stationnaire et causal. Celui-ci le sera si toutes les racines du polynôme  $\Theta(x)$  sont en dehors du cercle unité complexe. Autrement dit, il faut que  $\forall x \in \mathbb{C}$  tel que  $|x| \leq 1$ ,  $\Theta(x) \neq 0$ .

Pour estimer l’ordre  $p$  d’un modèle AR à partir d’une série donnée, on procède de la même façon que pour un modèle MA, mais en considérant la version partielle de la fonction

d'autocorrélation (pacf). En considérant le processus AR(3) de la figure 1.2 (graphique de gauche), on obtient la pacf illustrée à la même figure 1.2 (graphique de droite) qui semble indiquer qu'à partir du délai  $l = 4$  la fonction d'autocorrélation partielle est nulle, d'où  $p = 3$ . Encore ici, utiliser le AIC ou le SBC serait tout aussi approprié.

### 1.1.3 ARMA

À partir des modèles AR( $p$ ) et MA( $q$ ) introduits, on obtient le modèle ARMA (Auto-régressif Moyenne Mobile). On dit qu'un processus  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation d'un processus ARMA d'ordre  $p$  et  $q$ , noté ARMA( $p, q$ ), si et seulement si

$$\Theta(B) X_t = v + \Gamma(B) \epsilon_t, \quad (1.10)$$

où  $v \in \mathbb{R}$ ,  $\Theta(B)$  et  $\Gamma(B)$  sont les mêmes qu'en (1.9) et (1.7) respectivement et  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc gaussien. Par convention, on note que  $\theta_0 = \tau_0 = 1$  et que  $\theta_p, \tau_q \neq 0$ .

On observe que les processus AR( $p$ ) et MA( $q$ ) sont des cas particuliers du processus ARMA( $p, q$ ). En effet, le processus AR( $p$ ) correspond à un processus ARMA( $p, 0$ ) et le processus MA( $q$ ) correspond à ARMA( $0, q$ ).

Pour qu'un processus ARMA( $p, q$ ) soit stationnaire, comme pour un modèle AR, il faut que toutes les racines du polynôme  $\Theta(x)$  soient à l'extérieur du cercle unité complexe. Ensuite, ce même processus ARMA( $p, q$ ) est dit inversible, comme pour un modèle MA, si toutes les racines de  $\Gamma(x)$  sont à l'extérieur du cercle unité complexe.

En supposant la stationnarité du processus ARMA( $p, q$ ), on note aussi que

$$E[X_t] = \frac{v}{1 - \theta_1 - \dots - \theta_p}.$$

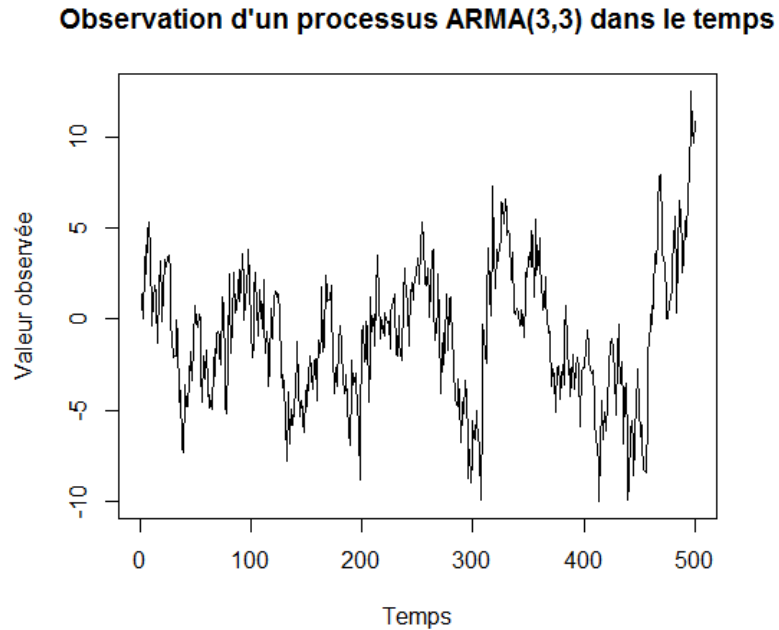
et que la fonction d'autocorrélation  $\rho(l)$  satisfait

$$\rho(l) - \theta_1 \rho(l-1) - \dots - \theta_p \rho(l-p) = 0 \quad \text{pour } l > q$$



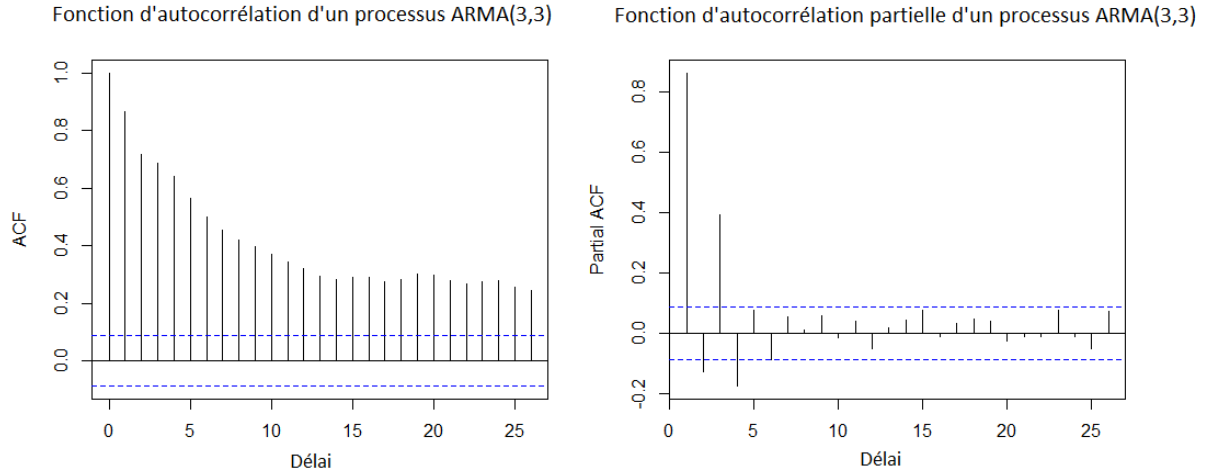
Pour trouver les ordres  $p$  et  $q$  d'un processus  $\text{ARMA}(p,q)$ , la double observation des fonctions d'autocorrélation et d'autocorrélation partielle est une façon de procéder, mais demeure un outil visuel et il n'est pas toujours évident d'en déduire les valeurs de  $p$  et de  $q$ . La difficulté liée à l'obtention des ordres d'un processus ARMA est présentée dans l'exemple aux figures 1.3 et 1.4 qui illustrent respectivement 500 observations d'un processus  $\text{ARMA}(3,3)$  avec  $\tau_1 = \theta_1 = 1$ ,  $\tau_2 = \theta_2 = -0,5$ ,  $\tau_3 = \theta_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $\mu = c = 0$  et les fonctions d'autocorrélation et d'autocorrélation partielle associées. Bref, l'utilisation du AIC [Aka74] ou du SBC [Sch78] est un meilleur choix.

Figure 1.3 – Graphique d'une réalisation partielle de 500 observations d'un processus  $\text{ARMA}(3,3)$  de paramètres  $\tau_1 = \theta_1 = 1$ ,  $\tau_2 = \theta_2 = -0,5$ ,  $\tau_3 = \theta_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $\mu = v = 0$



Le modèle ARMA, en plus de généraliser les processus autorégressif et moyenne mobile, permet d'éviter l'utilisation d'un processus AR ou d'un processus MA de grand ordre. Le tout a comme optique d'utiliser et d'estimer le moins de paramètres possible. L'esti-

Figure 1.4 – Graphiques des fonctions d'autocorrélation et d'autocorrélation partielle d'une réalisation partielle de 500 observations d'un processus ARMA(3,3) de paramètres  $\tau_1 = \theta_1 = 1$ ,  $\tau_2 = \theta_2 = -0,5$ ,  $\tau_3 = \theta_3 = 0.4$ ,  $\sigma_\epsilon = \sqrt{0,25}$  et  $\mu = v = 0$



mation de ces dits paramètres est effectuée généralement par des méthodes reposant sur le maximum de vraisemblance et ont été principalement élaborées par Box et Jenkins (1976) [BJ76]. Les ouvrages de Box et Jenkins (1976) [BJ76] et de Brockwell et Davis (1991) [BD91] sont d'excellentes références sur le sujet.

#### 1.1.4 VAR

Les séries variant dans le temps le font souvent en compagnie d'autres séries. Par exemple, un indice boursier varie avec un ensemble d'autres indices. L'intérêt de considérer un ensemble de séries chronologiques est d'étudier la structure de dépendance entre les séries et dans une série elle-même. Il existe, pour la plupart des modèles univariés, une généralisation aux cas multivariés. Dans ce travail, tout l'intérêt sera porté sur le modèle VAR (vecteur autorégressif) qui est une extension du modèle AR. C'est avec ce modèle classique multidimensionnel que des liens seront établis, au chapitre 3, avec la décomposition

de type M-Vine.

Ainsi, comme VAR est une extension multidimensionnelle du AR, sa représentation est la même sauf que les paramètres seront désormais des matrices réelles. On dit que le processus  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$   $k$ -varié satisfait une représentation d'un processus vecteur autorégressif d'ordre  $p$  (VAR( $p$ )) si et seulement si

$$\Theta(B) \mathbf{X}_t = \mathbf{v} + \boldsymbol{\epsilon}_t,$$

où  $\mathbf{v}$  est un vecteur réel  $k$ -varié,

$$\Theta(B) = \sum_{j=0}^p \boldsymbol{\theta}_j B^j,$$

$(\boldsymbol{\epsilon}_t)_{t \in \mathbb{Z}}$  est un vecteur bruit blanc gaussien de dimension  $k$  tel que  $\boldsymbol{\epsilon}_t \sim N(0, \Sigma)$ , et  $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p$  sont des matrices de réelles  $k \times k$  dont on a, par convention,  $\boldsymbol{\theta}_p \neq O$ , la matrice nulle, et  $\boldsymbol{\theta}_0 = \mathbb{I}$ , la matrice identité.

Dans un exemple avec  $k = 2$  et  $p = 1$ , on obtient :

$$\mathbf{X}_t = \mathbf{v} + \boldsymbol{\theta}_1 \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$$

où

$$\begin{aligned} \mathbf{X}_t &= (X_{1;t}, X_{2;t})^T, \\ \mathbf{v} &= (v_1, v_2)^T, \\ \boldsymbol{\theta}_1 &= \begin{pmatrix} \theta_{1,1}^{(1)} & \theta_{1,2}^{(1)} \\ \theta_{2,1}^{(1)} & \theta_{2,2}^{(1)} \end{pmatrix}, \\ \boldsymbol{\epsilon}_t &= (\epsilon_{1;t}, \epsilon_{2;t})^T \text{ et} \\ \Sigma &= \begin{pmatrix} E(X_{1;t}^2) & E(X_{1;t}X_{2;t}) \\ E(X_{1;t}X_{2;t}) & E(X_{2;t}^2) \end{pmatrix}. \end{aligned}$$

L'une des particularités des processus vecteur autorégressif est que  $\forall p \in \mathbb{N}^*$  un processus VAR( $p$ )  $k$ -varié peut s'écrire sous la forme d'un  $(p \times k)$ -varié VAR(1). En effet, soit

$(\mathbf{X}_t)_{t \in \mathbb{Z}}$  un processus VAR( $p$ )  $k$ -varié, alors il s'écrit sous la forme suivante :

$$\mathbf{X}_t = \mathbf{v} + \boldsymbol{\theta}_1 \mathbf{X}_{t-1} + \dots + \boldsymbol{\theta}_p \mathbf{X}_{t-p} + \boldsymbol{\epsilon}_t$$

d'où

$$\begin{pmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ O \\ \vdots \\ O \end{pmatrix} + \begin{pmatrix} \boldsymbol{\theta}_1 & \cdots & \boldsymbol{\theta}_{p-1} & \boldsymbol{\theta}_p \\ \mathbb{I} & \cdots & O & O \\ \vdots & \ddots & \vdots & \vdots \\ O & \cdots & \mathbb{I} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_{t-2} \\ \vdots \\ \mathbf{X}_{t-p} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_t \\ O \\ \vdots \\ O \end{pmatrix}.$$

On dit que le processus  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$  est stable si les racines de

$$\det(\mathbb{I} - \boldsymbol{\theta}_1 z - \dots - \boldsymbol{\theta}_p z^p) = 0$$

sont en dehors du cercle unité complexe, ou, de façon équivalente, si les valeurs propres de la matrice

$$\begin{pmatrix} \boldsymbol{\theta}_1 & \cdots & \boldsymbol{\theta}_{p-1} & \boldsymbol{\theta}_p \\ \mathbb{I} & \cdots & O & O \\ \vdots & \ddots & \vdots & \vdots \\ O & \cdots & \mathbb{I} & O \end{pmatrix}$$

sont de modules inférieurs à 1. La stabilité d'un processus VAR( $p$ ) implique la stationnarité de celui-ci.

L'estimation des paramètres d'un processus VAR( $p$ ) est généralement effectuée par la méthode des moindres carrés ordinaires (OSL). Canova (2017) [Can17] est une excellente référence pour l'estimation des paramètres d'un VAR( $p$ ). L'estimation par maximum de vraisemblance est aussi une méthode utilisée.

### 1.1.5 (G)ARCH

L'étude d'une série chronologique force parfois un chercheur à considérer une certaine volatilité pour ce processus. Dans les modèles univariés AR, MA et ARMA, mentionnés

dans la section ci-haut, la variance de ces processus est déterminée par la variance du processus des innovations et celle-ci est stable dans le temps. Ainsi, ces modèles ne sont pas utiles dans l'étude de séries pour lesquelles il y a des variations de la variance. Engle (1982) [Eng82] introduit un modèle pour les séries chronologiques qui prend en considération la volatilité de la variance dans le temps : le modèle ARCH (modèle autorégressif conditionnellement hétéroscédastique).

Engle [Eng82] mentionne qu'une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  satisfait une représentation de ARCH d'ordre  $q \geq 1$ , notée ARCH( $q$ ) si et seulement si

$$X_t = \sigma_t \epsilon_t \quad \text{et} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 \quad (1.11)$$

où  $\alpha_0, \alpha_1, \dots, \alpha_q \in \mathbb{R}^+$  et  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc tel que  $\epsilon_t$  est indépendant de  $X_{t-l}$   $\forall l \geq 1$ .

Avec ce modèle, la volatilité est expliquée par les  $q$  mesures de volatilité précédentes, c'est-à-dire pour les  $q$  innovations précédentes  $\epsilon_t$ . La figure 1.5, illustre un graphique d'un processus ARCH(2) avec  $\alpha_0 = 0$ ,  $\alpha_1 = 0.2$  et  $\alpha_2 = 0.4$ . La simple observation du graphique est suffisante pour constater la volatilité du processus.

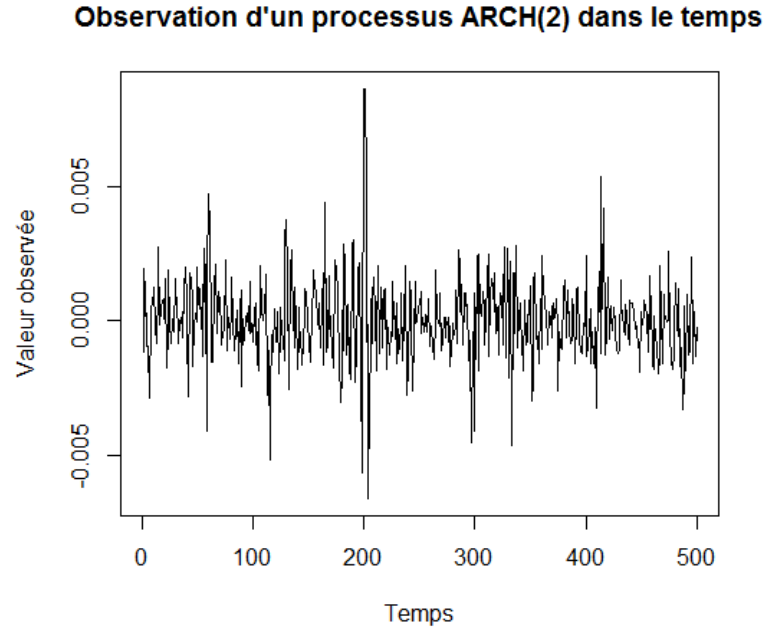
Une généralisation du modèle ARCH( $q$ ) a été proposée en 1986 par Bollerslev[Bol86]. Ce modèle ARCH généralisé est noté GARCH (modèle autorégressif conditionnellement hétéroscédastique généralisé).

Bollerslev [Bol86] mentionne qu'une série chronologique  $(X_t)_{t \in \mathbb{Z}}$  satisfait la représentation d'un modèle GARCH d'ordre  $p$  et  $q$ , noté GARCH( $p, q$ ), si et seulement si

$$X_t = \sigma_t \epsilon_t \quad \text{et} \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (1.12)$$

où  $\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p \in \mathbb{R}^+$  et  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc tel que  $\epsilon_t$  est indépendant de  $X_{t-l}$   $\forall l \geq 1$ .

Figure 1.5 – Graphique d’une réalisation partielle de 500 observations d’un processus ARCH(2) de paramètres  $\alpha_0 = 0$ ,  $\alpha_1 = 0.2$  et  $\alpha_2 = 0.4$



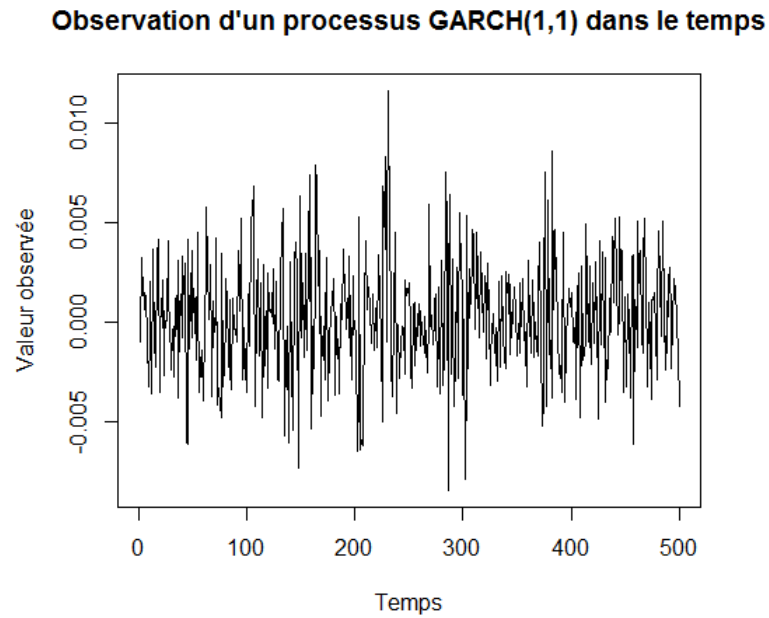
Le modèle GARCH généralise le modèle ARCH en permettant à la variance conditionnelle de suivre un processus ARMA( $p, q$ ) via l’introduction d’une composante autorégressive. La figure 1.6 illustre un processus GARCH(1,1) avec  $\alpha_0 = 0$ ,  $\alpha_1 = 0,2$  et  $\beta_1 = 0.7$ . On peut y observer la volatilité

Les modèles ARCH et GARCH, bien que décrits comme processus univariés, peuvent se généraliser au cas multivarié. Cependant, comme ces modèles ne sont pas les objets de cette étude, l’approfondissement des concepts et des idées reliées aux processus dit conditionnellement hétéroscédastique n’ira pas plus loin. On réfère le lecteur aux travaux de Engle [Eng82] et de Bollerslev [Bol86] pour plus de détails sur les modèles ARCH et GARCH respectivement.

D’autres modèles pour analyser les séries existent dans la littérature mais ceux-ci ne se-

ront pas considérés dans ce travail, on cite en exemple la volatilité stochastique. Gatheral (2006) [Gat06] aborde le sujet de la volatilité stochastique.

Figure 1.6 – Graphique d’une réalisation partielle de 500 observations d’un processus GARCH(1,1) de paramètres  $\alpha_0 = 0$ ,  $\alpha_1 = 0.2$  et  $\beta_1 = 0.7$



# CHAPITRE 2

## Copules

Le concept de dépendance en statistique ne date pas d'hier et plusieurs mesures ont été développées pour quantifier cette dépendance. La mesure la plus connue est certainement le coefficient de corrélation linéaire de Pearson, noté  $r$ . Soient  $X$  et  $Y$  deux variables aléatoires dont on observe  $n$  réalisations  $x_1, x_2, \dots, x_n$  et  $y_1, y_2, \dots, y_n$ . On définit le coefficient de corrélation  $r$  de Pearson empirique par :

$$r = \frac{\widehat{Cov}(X, Y)}{s_X s_Y}$$

où  $s_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$  et  $s_Y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$  sont les écart-types des échantillons tirés de  $X$  et  $Y$  respectivement et  $\widehat{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$  est la covariance échantillonnale de  $X$  et  $Y$ .

Des mesures comme le  $\tau$  de Kendall et le coefficient de corrélation de rang de Spearman sont d'autres outils permettant de quantifier la dépendance entre des variables aléatoires à partir d'un jeu de données. En notant  $R_{i,n}^X = \{rg(x_1), rg(x_2), \dots, rg(x_n)\}$  et  $R_{i,n}^Y = \{rg(y_1), rg(y_2), \dots, rg(y_n)\}$  les ensembles des rangs des valeurs des variables aléatoires



respectives  $X$  et  $Y$ , on a

$$\tau_{Kendall} = \frac{(\text{Nombre de paires concordantes}) - (\text{Nombre de paires discordantes})}{\frac{1}{2}n(n-1)}$$

$$r_{Spearman} = \frac{\widehat{Cov}(rg(X), rg(Y))}{s_{rg(X)}s_{rg(Y)}}$$

où on note que  $\widehat{Cov}(rg(X), rg(Y)) = \frac{1}{n} \sum_{i=1}^n (rg(x_i) - \overline{rg(x)}) (rg(y_i) - \overline{rg(y)})$ ,  $s_{rg(X)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (rg(x_i) - \overline{rg(x)})^2}$ ,  $s_{rg(Y)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (rg(y_i) - \overline{rg(y)})^2}$ ,  $\overline{rg(x)} = \frac{1}{n} \sum_{i=1}^n rg(x_i)$  et  $\overline{rg(y)} = \frac{1}{n} \sum_{i=1}^n rg(y_i)$ . De plus, prendre note qu'avec un ensemble de réalisations  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  provenant de variables aléatoires jointes  $X$  et  $Y$ , on a que des paires d'observations  $(x_i, y_i)$  et  $(x_j, y_j)$  sont dites concordantes si  $x_i < x_j$  et  $y_i < y_j$  ou si  $x_i > x_j$  et  $y_i > y_j$ .

Le principal défaut de ces mesures de dépendance est qu'elles sont d'une flexibilité limitée et décrivent la dépendance par une seule valeur ou une seule matrice. Une alternative intéressante à ce problème de non flexibilité est l'utilisation de copules. Celles-ci permettent à la dépendance de varier selon les valeurs prises par la variable aléatoire. En effet, il devient possible de représenter la dépendance entre deux variables aléatoires par une fonction de répartition.

Dans ce chapitre, c'est justement le concept de copules ainsi que leurs propriétés qui sont abordés. Dans le cadre de ce mémoire, on s'intéressera aux copules pour étudier la structure de dépendance de séries chronologiques et particulièrement au cas multidimensionnel pour lequel la décomposition de type Vine sera présentée et discutée en détail.

Dans la section 3.1 de ce chapitre, la définition et différentes propriétés des copules sont introduites. Ensuite, dans la section 3.2, différentes familles de copules sont présentées. Finalement, dans la section 3.3, la décomposition de type Vine des copules multivariées est abordée : d'abord les décompositions C-Vine et D-Vine et ensuite la décomposition M-Vine introduite par Beare et Seo (2015) [BS15] qui est spécifique à l'étude des structures

de dépendance pour les séries chronologiques.

## 2.1 Définition et propriétés

Comme mentionné plus haut, la copule est un outil qui permet d'étudier la structure de dépendance entre deux ou plusieurs variables aléatoires. Elle permet de modéliser la structure de dépendance et de construire des fonctions de répartition. Donc, au lieu de représenter la structure de dépendance entre deux séries par une valeur de corrélation, il devient possible d'obtenir une fonction de densité qui permet d'étudier la dépendance de façon ponctuelle et de façon plus informative.

C'est Sklar [Skl59] qui définit, en 1959, les copules à l'aide d'un théorème qui stipule :

*Soient  $F$  une fonction de répartition conjointe  $k$ -variée continue de  $(X_1, X_2, \dots, X_k)$  et  $F_i$  la fonction de répartition marginale continue de  $X_i$  pour  $i = 1, 2, \dots, k$ , alors il existe une unique fonction de répartition  $k$ -variée  $C : [0, 1]^k \rightarrow [0, 1]$ , appelée copule, avec marges uniformes telle que  $\forall (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ , on a :*

$$F(x_1, x_2, \dots, x_k) = C(F_1(x_1), F_2(x_2), \dots, F_k(x_k)). \quad (2.1)$$

Ce théorème est aussi la première définition formelle de copule. Le concept n'est pas, ou très peu utilisé, durant les années suivant son introduction. Ce n'est qu'à partir de la décennie 90 que les copules se font voir régulièrement dans la littérature mathématique avec notamment des livres de Joe (1997) [Joe97] et Nelsen (1999) [Nel99]. Avec ce concept de copules, si  $F_j$  est continue  $\forall j \in \{1, 2, \dots, n\}$ , on a directement de (2.1) que pour  $(u_1, u_2, \dots, u_k) \in [0, 1]^k$ ,

$$C(u_1, u_2, \dots, u_k) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_k^{-1}(u_k)), \quad (2.2)$$

où  $F_j^{-1}$  est un inverse généralisée de  $F_j$ . L'expression (2.2) permet de définir une copule

à partir d'une distribution conjointe de répartition et des distributions marginales de répartition inverses. Cette formule permet d'estimer la copule  $C$ . En effet, l'estimation de copules est basée sur cette équation.

De plus, à partir de cette même expression (2.2), il devient possible de définir la densité d'une copule lorsqu'elle existe, notée  $c(u_1, u_2, \dots, u_k)$ , à partir de  $C(u_1, u_2, \dots, u_k)$  comme suit :

$$c(u_1, u_2, \dots, u_k) = \frac{\partial^k}{\partial u_1 \partial u_2 \dots \partial u_k} C(u_1, u_2, \dots, u_k). \quad (2.3)$$

Cette expression correspond à une fonction de densité pour la structure de dépendance de variables aléatoires définie sur  $[0, 1]^k$ . Effectivement, il découle de (2.3) que la densité d'une distribution conjointe de  $(X_1, X_2, \dots, X_k)$ , notée  $f$ , peut être écrite à l'aide des densités et des répartitions marginales de  $X_i$ , notée respectivement  $f_i$  et  $F_i$  pour  $i = 1, 2, \dots, k$ , et de densité de copule  $k$ -variée  $c$  comme suit :

$$f(x_1, x_2, \dots, x_k) = c(F_1(x_1), F_2(x_2), \dots, F_k(x_k)) \times f_1(x_1) f_2(x_2) \dots f_k(x_k) \quad (2.4)$$

Ainsi, la densité conjointe d'une distribution  $k$ -variée peut s'écrire comme le produit des densités marginales et d'une copule de densité  $k$ -variée. Donc, cette dernière densité de copule décrit à elle seule la structure de dépendance entre les variables aléatoires  $X_1, X_2, \dots, X_k$ . Cette dernière expression sert à l'estimation puisqu'elle définit la fonction de vraisemblance.

## 2.2 Classes et familles de copules

Dans un esprit paramétrique de l'étude des copules, plusieurs familles de copules ont été introduites au cours des années. Un grand nombre de familles a déjà été proposé et étudié par divers chercheurs et auteurs. Plusieurs familles de copules partageant des similarités sont aussi regroupées en classes. Les classes les plus discutées sont les classes elliptique et archimédienne. Dans cette section, la description des classes elliptique et archimédienne est faite suivie de la présentation de quelques unes des familles les plus utilisées, c'est-à-dire les familles de copule d'indépendance, gaussienne, de Student, de Clayton, de Frank, de Gumbel et d'Ali-Mikhail-Haq.

### 2.2.1 Classe elliptique

Une copule  $C$  est dite elliptique si et seulement si la densité  $c$  de cette dite copule peut s'écrire sous la forme suivante :

$$c(\mathbf{u}) = (\det(\Sigma))^{-1/2} g\left((\mathbf{u} - \boldsymbol{\mu}) \Sigma^{-1} (\mathbf{u} - \boldsymbol{\mu})^T\right),$$

où  $\mathbf{u} \in [0, 1]^k$ ,  $\boldsymbol{\mu} \in \mathbb{R}^k$  est un paramètre de position,  $\Sigma$  est une matrice symétrique et définie positive et  $g(\cdot)$  est une fonction à valeurs positives vérifiant  $\int_{\mathbb{R}^k} g(\mathbf{x}\mathbf{x}^T) d\mathbf{x} = 1$ .

Dans le cas normal,  $\boldsymbol{\mu}$  est la moyenne d'une loi multivariée et  $\Sigma$  en est la matrice de variance-covariance. Quant à  $g(\cdot)$ , elle devient la fonction de densité de la loi multivariée sous-jacente.

La classe elliptique est un regroupement de familles de copules ayant cette particularité. On y notera notamment la famille de copules gaussiennes (2.2.4) et de Student 2.2.5).

### 2.2.2 Classe archimédienne

La particularité des copules archimédiennes est que celles-ci sont décrites à l'aide d'une fonction génératrice  $\varphi : [0, 1] \rightarrow [0, \infty)$  ayant comme propriété que  $\forall u \in [0, 1] \varphi(1) = 0$ ,  $\frac{d\varphi}{du}(u) < 0$  et  $\frac{d^2\varphi}{du^2}(u) > 0$ .

Alors l'expression d'une copule archimédienne est :

$$C(u_1, u_2, \dots, u_k) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_k)) & \text{si } \varphi(u_1) + \dots + \varphi(u_k) \leq \varphi(0) \\ 0 & \text{sinon.} \end{cases}$$

Donc, toute copule archimédienne est définie par sa fonction génératrice. Les copules de Clayton, de Frank, de Gumbel et d'ali-Mikhail-Haq, font parties de la famille des copules archimédiennes. Avec ces deux dernières classes de copules définies, on s'intéresse désormais à quelques familles de copules.

### 2.2.3 Copule d'indépendance

La copule d'indépendance, comme son nom l'indique, représente l'indépendance entre les variables aléatoires. Donc, si  $X_1, X_2, \dots, X_k$  sont des variables aléatoires indépendantes, on a que

$$f(x_1, x_2, \dots, x_k) = f_1(x_1) f_2(x_2) \dots f_k(x_k)$$

où  $f$  est la densité conjointe de  $(X_1, \dots, X_k)$  et  $f_i$  est la densité marginale de  $X_i$  pour  $i = 1, 2, \dots, k$ .

En se basant sur l'expression (2.4), on déduit que  $c(F_1(x_1), F_2(x_2), \dots, F_k(x_k)) = 1$  représente l'indépendance entre les variables. Autrement dit, pour une copule d'indépendance, on a que  $\forall (u_1, \dots, u_k) \in [0, 1]^k$

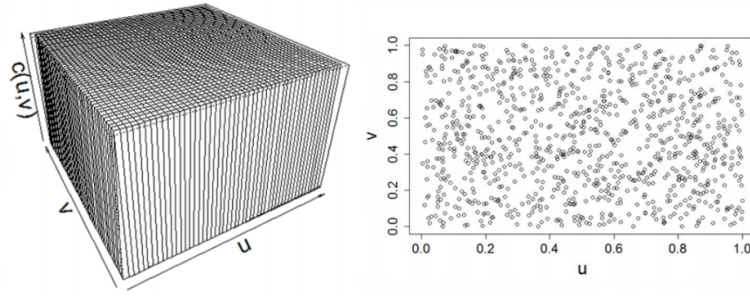
$$c(u_1, u_2, \dots, u_k) = 1$$

d'où

$$C(u_1, u_2, \dots, u_k) = \int_0^{u_1} \int_0^{u_2} \dots \int_0^{u_k} 1 dt_1 dt_2 \dots dt_k = u_1 u_2 \dots u_k.$$

La figure 2.1 illustre la fonction de densité d'une copule d'indépendance bivariée ainsi qu'un nuage de points obtenu à partir de cette dernière.

Figure 2.1 – Graphique de densité et nuage de points d'une copule d'indépendance bivariée  $C(u, v)$



## 2.2.4 Copule gaussienne

Soient  $\Phi$  la fonction de répartition d'une distribution normale univariée centrée et réduite et  $\Phi_\Omega$  la fonction de répartition d'une distribution normale  $k$ -variée centrée et réduite avec

$$\Omega = \begin{pmatrix} 1 & \rho_{1;2} & \rho_{1;3} & \dots & \rho_{1;k} \\ \rho_{1;2} & 1 & \rho_{2;3} & \dots & \rho_{2;k} \\ \vdots & & \ddots & & \vdots \\ \rho_{1;k} & \rho_{2;k} & \rho_{3;k} & \dots & 1 \end{pmatrix}$$

la matrice des corrélations entre les  $k$  variables aléatoires  $X_1, X_2, \dots, X_k$ . Alors, la copule gaussienne est définie comme suit :

$$C(u_1, \dots, u_k; \Omega) = \Phi_\Omega(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_k)).$$

Cette formule découle directement de l'équation (2.2). On peut la réécrire, en élaborant, comme suit :

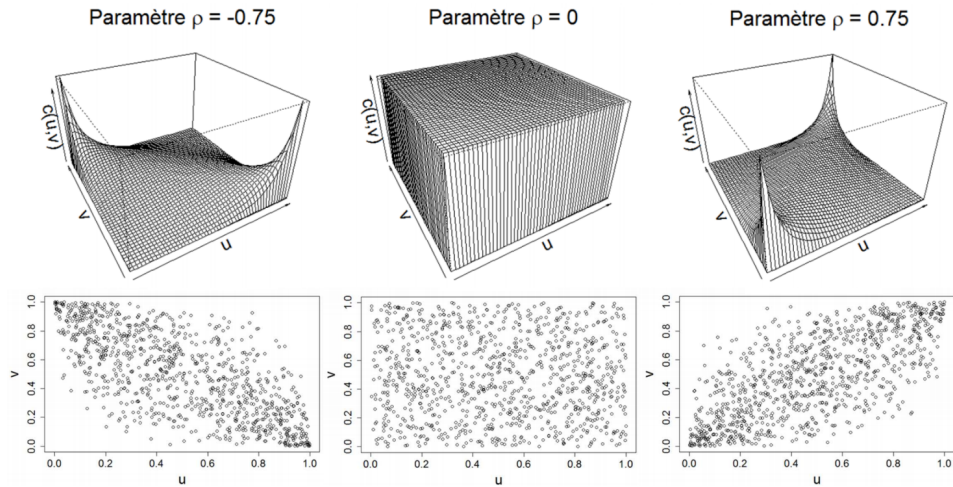
$$c(u_1, \dots, u_k; \Omega) = \frac{1}{\sqrt{\det(\Omega)}} \exp \left( -\frac{1}{2} \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_k) \end{pmatrix}^T (\Omega^{-1} - \mathbb{I}) \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_k) \end{pmatrix} \right)$$

qui correspond à la définition d'une copule elliptique. Donc, la famille des copules gaussiennes fait partie de la classe des copules elliptiques.

On note que dans le cas où  $\Omega = I$ , on obtient que  $C(u_1, \dots, u_k; \Omega) = u_1 \dots u_k$  qui est la copule d'indépendance. La figure 2.2 illustre des densités de copules gaussiennes bivariées et des nuages de points associés avec différentes valeurs de  $\rho$  (puisque'il est question du cas bivarié, l'ensemble  $\Omega$  ne comporte qu'une corrélation notée  $\rho$ ).

Cette famille de copules ne décrit que les relations linéaires entre les variables aléatoires et elle ne présente donc pas les dépendances aux extrêmes.

Figure 2.2 – Graphiques de densité et nuages de points de copules gaussiennes bivariées  $C(u, v; \rho)$  avec  $\rho = -0,75$ ,  $\rho = 0$  et  $\rho = 0,75$



## 2.2.5 Copule de Student

Soient  $T_\kappa$  la fonction de répartition de la distribution de Student avec  $\kappa \geq 3$  degrés de liberté et  $T_{\Omega,\kappa}$  la distribution de Student multivariée de degré de liberté  $\kappa \geq 3$  avec comme matrice de forme  $\Omega$ , l'ensemble des corrélations entre les variables aléatoires. Alors, la copule de Student est définie comme suit :

$$C(u_1, \dots, u_n; \Omega, \kappa) = T_{\Omega,\kappa}(T_\kappa^{-1}(u_1), \dots, T_\kappa^{-1}(u_n)).$$

Cette dernière expression peut être élaborée pour obtenir la densité de copule  $c$  suivante :

$$c(u_1, \dots, u_n; \Omega, \kappa) = \frac{t_{\Omega,\kappa}(T_\kappa^{-1}(u_1), \dots, T_\kappa^{-1}(u_n))}{\prod_{i=1}^n f_\kappa(T_\kappa^{-1}(u_i))}$$

où  $T_\kappa^{-1}$  est la fonction de répartition inverse de  $T_\kappa$ ,  $t_{\Omega,\kappa}$  est la fonction de densité  $T_{\Omega,\kappa}$  et  $f_\kappa$  est la densité univariée d'une loi de Student avec paramètre de corrélation  $\rho = 1$ .

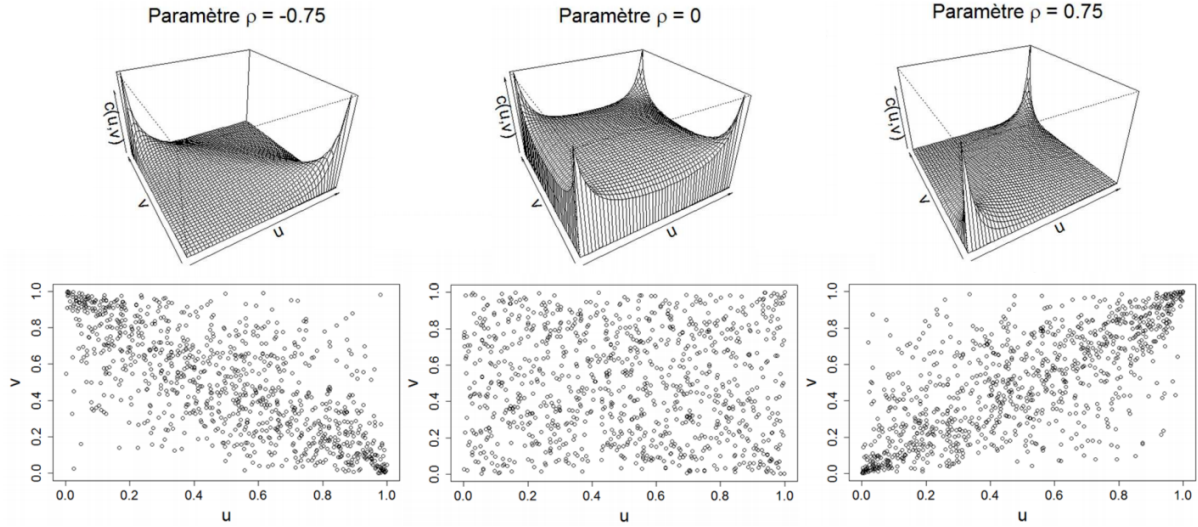
Les copules de Student, tout comme les copules gaussiennes, sont dites de la famille des copules elliptiques. L'avantage de la copule de Student vis-à-vis la copule gaussienne est que cette première est capable de détecter une forme de dépendance dans les queues (aux extrêmes).

Pour illustrer la famille des copules de Student, on considère dans la figure 2.3 des densités de copules bivariées de Student de degré de liberté  $\kappa = 5$  et de corrélations  $\rho = -0,75$ ,  $\rho = 0$  et  $\rho = 0,75$ . Pour chacune des copules de Student considérées, on illustre aussi un nuage de points issu de la copule.

On y constate que même avec un paramètre  $\rho = 0$ , on n'obtient pas une copule d'indépendance. En effet, il y a une concentration plus élevée de points dans les extrêmes.



Figure 2.3 – Graphiques de densité et nuages de points de copules de Student bivariées  $C(u, v; \rho, \kappa = 5)$  avec  $\rho = -0,75$ ,  $\rho = 0$  et  $\rho = 0,75$



## 2.2.6 Copule de Clayton

La copule de Clayton est une copule de la classe archimédienne définie par la fonction génératrice  $\varphi_\theta(u) = \frac{1}{\theta}(u^{-\theta} - 1)$  avec  $\theta \in [-1, \infty)$  et  $\theta \neq 0$  avec laquelle on obtient la copule suivante :

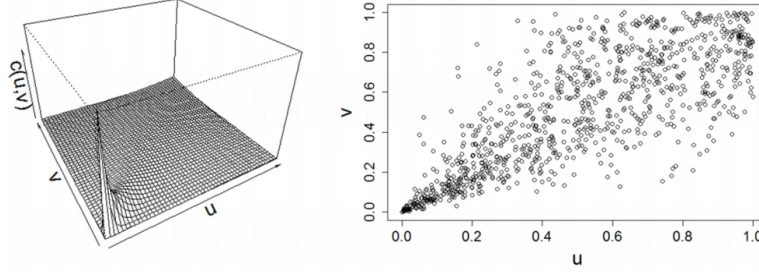
$$C_\theta(u_1, u_2, \dots, u_k) = [\max(u_1^{-\theta} + u_2^{-\theta} + \dots + u_k^{-\theta} - k + 1, 0)]^{-1/\theta}.$$

La copule de Clayton présente la caractéristique d'asymétrie suivante : elle ne détecte pas de dépendance pour de grandes valeurs, mais seulement pour les petites valeurs. La figure 2.4 illustre la densité d'une copule de Clayton de paramètre  $\theta = 3$  avec un nuage de 500 points simulés à partir de celle-ci. On y observe une accumulation concentrée de point pour des petites valeurs (près de (0,0)) qui implique une grande dépendance et aucune concentration comparable pour de grandes valeurs (près de (1,1)) qui implique une faible dépendance.

Il est possible de montrer que le tau de Kendall pour une copule de Clayton de fonction

génératrice  $\varphi_\theta(u)$  telle qu'exprimée plus haut est donné par  $\tau = \frac{\theta}{\theta+2}$  [GR93].

Figure 2.4 – Graphique de densité et nuage de points d'une copule de Clayton bivariée  $C_\theta(u, v)$  avec  $\theta = 3$



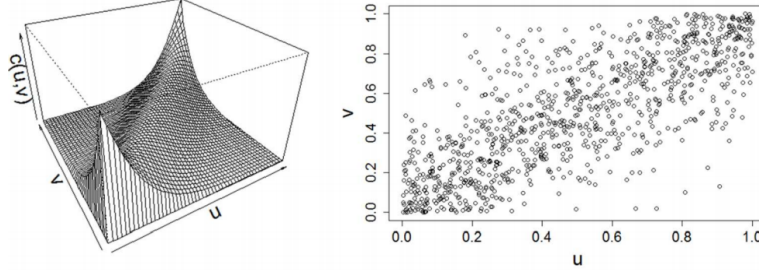
### 2.2.7 Copule de Frank

La copule de Frank fait partie de la famille archimédienne et a pour générateur la fonction de paramètre  $\theta$  :  $\varphi_\theta(u) = -\ln\left(\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right)$  avec  $\theta \in \mathbb{R}^*$ . Ainsi, la copule de Frank  $k$ -variée s'écrit :

$$C_\theta(u_1, u_2, \dots, u_k) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1) \dots (e^{-\theta u_k} - 1)}{(e^{-\theta} - 1)^{k-1}} \right).$$

Contrairement à la copule de Clayton, la copule de Frank est symétrique dans le sens que la dépendance dans les extrêmes (0,0) et (1,1) est la même. Le densité de la copule de Frank bivariée de paramètre  $\theta = 7$  et un nuage de points simulé à partir de celle-ci sont illustrés à la figure 2.5.

Figure 2.5 – Graphique de densité et nuage de points d’une copule de Frank bivariable  $C_\theta(u, v)$  avec  $\theta = 7$



### 2.2.8 Copule de Gumbel

La copule de Gumbel a pour fonction génératrice  $\varphi_\theta(t) = (-\ln t)^\theta$  avec  $\theta \geq 1$  d’où on a :

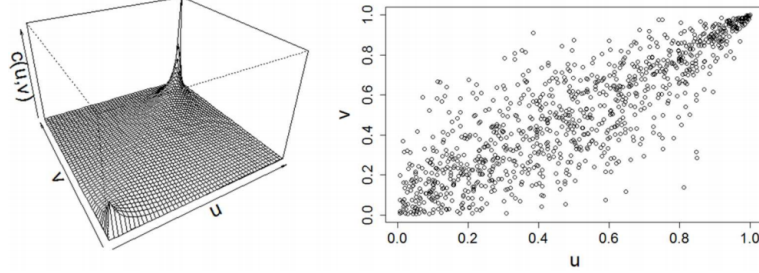
$$C_\theta(u_1, \dots, u_k) = e^{-[(-\ln u_1)^\theta + (-\ln u_2)^\theta + \dots + (-\ln u_k)^\theta]^{1/\theta}}.$$

Dans un cas bivariable avec  $\theta = 3$  on obtient la densité de copule de Gumbel et un nuage de points associé illustré à la figure 2.6.

Une particularité de la copule du Gumbel est que celle-ci présente des dépendances pour les extrêmes (0,0) et (1,1) mais d’importance significativement différente. En effet, sur la figure 2.6, on constate une très forte concentration de points pour l’extrême (1,1) et une concentration relativement forte de points pour l’extrême (0,0).

Il est intéressant de noter que le tau de Kendall pour la copule de Gumbel est donné par  $\tau = 1 - \frac{1}{\theta}$  [GR93].

Figure 2.6 – Graphique de densité et nuage de points d’une copule de Gumbel bivariée  $C_\theta(u, v)$  avec  $\theta = 3$



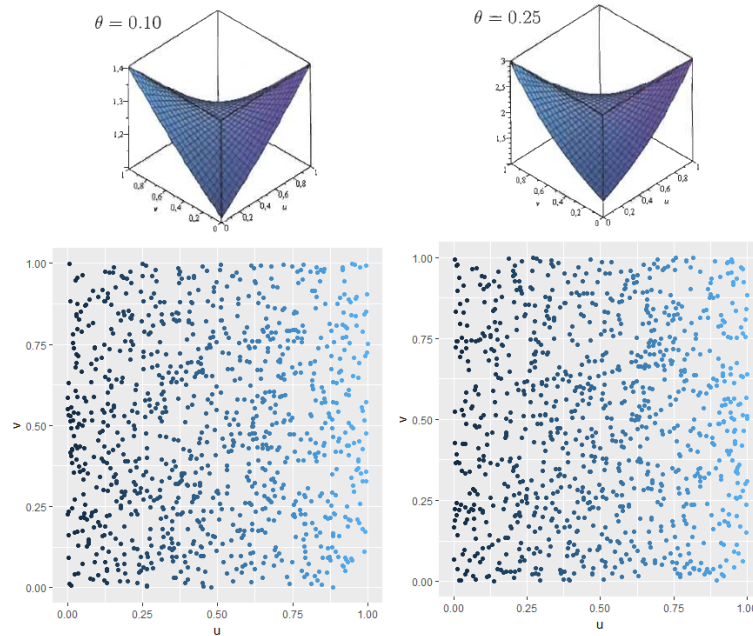
### 2.2.9 Copule d’Ali-Mikhail-Haq (AMH)

Finalement, la copule AMH a pour fonction génératrice  $\varphi_\theta(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right)$  avec  $\theta \in [-1, 1)$ . On obtient alors que la copule AMH s’écrit :

$$C_\theta(u_1, \dots, u_k) = \frac{u_1 u_2 \dots u_k}{1 - \theta(1 - u_1)(1 - u_2) \dots (1 - u_k)}.$$

On note que la copule AMH est la seule parmi les 22 copules archimédiennes étudiées dans l’article de Nelsen (1999) [Nel99] dont le paramètre  $\theta$  est dans un intervalle  $[0, 1]$  et qui mesure autant les dépendances positives que négatives. Aussi, la copule AMH montre une forte dépendance de queue à gauche (près de  $(0,0)$ ) quand le paramètre  $\theta = 1$ . De plus, lorsque le paramètre  $\theta$  tend vers 0, la copule AMH tend vers la copule d’indépendance. En bref, même en faisant partie des copules archimédiennes, cette copule AMH possède certaines particularités que possède aussi la copule gaussienne. La figure 2.7 illustre des exemples de densités avec nuages de points associés pour des copules AMH de paramètre respectif  $\theta = 0.10$  et  $\theta = 0.25$ .

Figure 2.7 – Graphiques de densité et nuages de points de copules d’Ali-Mikhail-Haq bivariées  $C_\theta(u, v)$  avec  $\theta = 0.10$  et  $\theta = 0.25$



## 2.3 Méthode d’estimation des copules

L’objectif pratique des copules étant de modéliser la dépendance entre des variables, il est crucial de connaître des méthodes d’estimation pour celles-ci. Cependant, ce sujet n’étant pas la finalité de ce travail, une présentation sommaire de différentes méthodes d’estimation y est faite.

Il existe principalement trois approches pour l’estimation de copules : paramétrique, semi-paramétrique et non-paramétrique. Les différences entre ces trois approches se rapportent essentiellement à la façon d’estimer les distributions marginales et puis les copules elles-mêmes. En effet, pour l’estimation de copules, on se sert de l’équation (2.2) :

$$C(u_1, u_2, \dots, u_k) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_k^{-1}(u_k)) .$$

Ainsi, il est nécessaire d’avoir accès aux marginales (ou à une estimation des marginales)

pour obtenir une estimation de la copule. Les prochaines sous-sections s'intéressent respectivement à ces trois approches.

### 2.3.1 Estimation paramétrique

Comme son nom l'indique, l'approche paramétrique consiste à estimer les distributions marginales des variables aléatoires ainsi que la copule associée à la structure de dépendance de façon paramétrique. La méthode paramétrique la plus commune, étudiée dans [Joe97] et [Joe05], est celle de la méthode du maximum de vraisemblance complète (*Full Maximum Likelihood*). On utilise cette méthode pour estimer conjointement les paramètres pour les marginales et la copule par maximum de vraisemblance. Un utilisateur devra donc considérer tous les arrangements de types de copules et de lois marginales possibles et les optimiser par maximum de vraisemblance. Ensuite, on sélectionne le modèle qui concorde avec l'arrangement qui a atteint le plus haut maximum de vraisemblance. Un inconvénient majeur de cette méthode est qu'elle peut s'avérer computationnellement chargée surtout pour de grandes dimensions et peut s'avérer longue à optimiser.

Pour pallier à ce problème, une méthode alternative est discutée dans [JX96]. Celle-ci consiste principalement à estimer les paramètres des marginales dans une première étape et ensuite, en considérant que les paramètres estimés pour les marginales sont optimaux, à estimer les paramètres de la copule. Les estimations aux deux étapes se font aussi par maximum de vraisemblance. Cette méthode, qui est nommé fonction d'inférence pour les marges, nécessite beaucoup moins de calculs, mais ne garantit pas l'optimalité de la solution. Le souci relié à l'optimalité de cette méthode est qu'il est possible que les marginales estimées dans la première étape ne mènent pas nécessairement au modèle paramétrique optimal lors de la deuxième étape, contrairement à la première méthode qui mène à coup sûr au meilleur modèle purement paramétrique. En pratique, la méthode de

la fonction d'inférence pour les marges est beaucoup plus utilisée que celle du maximum de vraisemblance complète principalement pour sa rapidité d'exécution et du fait que sa solution ne s'éloigne pas beaucoup, de façon générale, de la solution optimale.

Le désavantage principal de ces deux dernières méthodes d'estimation paramétrique réside justement dans l'aspect paramétrique. En voulant estimer des paramètres pour les marginales et les copules associés à des modèles prédéfinis on se restreint énormément. C'est pourquoi il est peut être préférable d'opter pour des approches semi-paramétriques, voir non-paramétriques.

### 2.3.2 Estimation semi-paramétrique

L'approche semi-paramétrique pour l'estimation de copule consiste essentiellement à évaluer les marginales et la copule de façon différente : l'une avec une méthode paramétrique et l'autre avec une méthode non-paramétrique. La procédure préconisée par les statisticiens est d'estimer en premier lieu les marginales de façon non-paramétrique à l'aide de la distribution empirique, puis d'estimer la copule paramétriquement avec la méthode du maximum de vraisemblance en se basant sur les estimations non-paramétriques des marginales. Il s'agit de la méthode CML (*Canonical Maximum Likelihood*) étudiée entre autres par Genest, Ghoudi et Rivest (1995) [GGR95] et Kim (2007) [KSS07].

En résumé, la méthode CML est identique à la méthode de la fonction d'inférence pour les marges sauf qu'on estime les marginales de façon non-paramétrique.

### 2.3.3 Estimation non-paramétrique

L'estimation purement non-paramétrique de copules est un sujet discuté par Chen et Huang (2007) [CH07]. Pour cette méthode, les marginales et la copule sont estimées sans

paramètre. Pour y arriver, il est nécessaire d'utiliser la copule empirique introduite pour la première fois par Deheuvels (1980) [Deh80] sous l'appellation de fonction empirique de dépendance. Par la suite, Fermanian (2002) [FRW02] a proposé une version lisse pour la copule empirique.

Il existe quelques méthodes pour effectuer une estimation non-paramétrique d'une copule. La plus utilisée est la méthode à noyau qui connaît plusieurs dérivées et variations. L'une de ces variations est basée sur la transformation des données et est proposée par Gijbels (2012) [GOV12]. Une autre définie par Chen et Huang (2005) [?] propose d'utiliser un noyau local linéaire pour contrer les problèmes de biais aux bornes. On réfère le lecteur aux précédents articles pour de amples informations sur les méthodes d'estimation non-paramétriques pour les copules.

## 2.4 Décomposition de type Vine

Un souci persiste quant à l'utilisation de copules pour étudier la structure de dépendance lorsqu'on travaille dans des situations  $k$ -variées avec  $k > 2$  : l'estimation d'une copule  $k$ -variée restreint le chercheur à utiliser une seule famille de copules en utilisant une approche paramétrique. Prenons l'exemple d'une copule de Frank, celle-ci s'écrit  $C_\theta(u_1, u_2, \dots, u_k) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1) \dots (e^{-\theta u_k} - 1)}{(e^{-\theta} - 1)^{k-1}} \right)$  et donc un seul paramètre contrôle la structure de dépendance des  $k$  variables aléatoires. Ainsi, estimer une structure de dépendance d'une variable aléatoire  $k$ -variée à l'aide d'une copule de Frank peut s'avérer biaisée si la structure de dépendance de la variable aléatoire ne correspond pas exactement à une copule de Frank.

Une solution à ce problème est la décomposition de type Vine de copules multivariées en copules bivariées. L'approche basée sur la décomposition de type Vine consiste à



considérer une densité de copule  $k$ -variée comme un produit de  $k(k-1)/2$  densités de copules bivariées. Ainsi, dans un cadre empirique, on peut estimer chaque structure de dépendance associée à chacune des copules bivariées par des copules différentes ; la copule  $k$ -variée initiale devient le produit de chacune de ces copules. Cette façon de procéder permet une plus grande flexibilité lors du choix du modèle à utiliser.

Un aspect intéressant par rapport à la décomposition de type Vine : celle-ci n'est pas unique. Ainsi, plusieurs décompositions en copules bivariées sont possibles pour une même copule  $k$ -variée. Dans cette présente section, trois décompositions de type Vine seront présentées : le C-Vine, le D-Vine et le M-Vine.

### 2.4.1 C-Vine et D-Vine

Avant d'aborder les concepts de C-Vine et de D-Vine, l'idée derrière la décomposition de type Vine doit être présentée, car tout découle du conditionnement. Tout d'abord, l'expression (2.4) indique que la densité d'une distribution conjointe est égale au produit d'une copule de densité multivariée et de chacune des densités marginales. Ensuite, on sait que

$$f_{1,\dots,i|i+1,\dots,k}(x_1, x_2, \dots, x_i | x_{i+1}, \dots, x_k) = \frac{f(x_1, x_2, \dots, x_k)}{g(x_{i+1}, \dots, x_k)} \quad (2.5)$$

où  $f_{1,\dots,i|i+1,\dots,k}$  est la densité conditionnelle de  $(X_1, \dots, X_i)$  étant donné  $(X_{i+1} = x_{i+1}, \dots, X_k = x_k)$ ,  $f$  est la densité de la distribution conjointe du vecteur  $(X_1, \dots, X_k)$  et  $g(x_{i+1}, \dots, x_k) = \int \int \dots \int f(x_1, \dots, x_k) dx_1 \dots dx_{i-1} dx_i$  la densité conjointe de  $(X_{i+1}, \dots, X_k)$ .

Notons aussi  $c_{ij|D}$  la densité de copule associée à la structure de dépendance entre les variables  $X_i$  et  $X_j$ ,  $F_{i|D}$  la fonction de répartition de la variable aléatoire  $X_i$  et  $f_{i|D}$  la densité de cette même variable  $X_i$ , toutes trois conditionnelles aux variables  $X_d$  avec  $d \in D$  où  $D$  est un ensemble prédéterminé d'indices avec  $i, j = 1, \dots, k$ .

Pour commencer, on considère un cas bivarié. Les densités conjointe et conditionnelle sont respectivement données par

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2) \quad \text{et} \quad f_{2|1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

d'où

$$f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2)$$

En augmentant d'une dimension, on peut obtenir :

$$\begin{aligned} f(x_1, x_2, x_3) &= f_{3|12}(x_3|x_1, x_2) f_{2|1}(x_2|x_1) f_1(x_1) \\ f_{2|1}(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2) \\ f_{3|12}(x_3|x_1, x_2) &= c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)|x_2) f_{3|2}(x_3|x_2) \\ f_{3|2}(x_3|x_2) &= c_{23}(F_2(x_2), F_3(x_3)) f_3(x_3), \end{aligned}$$

d'où on a :

$$\begin{aligned} f(x_1, x_2, x_3) &= f_1(x_1) f_2(x_2) f_3(x_3) \\ &\times c_{12}(F_1(x_1), F_2(x_2)) c_{23}(F_2(x_2), F_3(x_3)) \\ &\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)|x_2). \end{aligned}$$

Comme  $f(x_1, x_2, x_3) = c(F_1(x_1), F_2(x_2), F_3(x_3)) f_1(x_1) f_2(x_2) f_3(x_3)$ , où  $c$  est la densité de copule pour la structure de dépendance de la distribution conjointe tri-variée  $(X_1, X_2, X_3)$ , on a alors que :

$$\begin{aligned} c(F_1(x_1), F_2(x_2), F_3(x_3)) &= c_{12}(F_1(x_1), F_2(x_2)) c_{23}(F_2(x_2), F_3(x_3)) \quad (2.6) \\ &\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)|x_2) \end{aligned}$$

Donc, une densité de copule tri-variée peut s'écrire comme le produit de 3 copules bivariées. On appelle la séparation d'une copule  $k$ -variée en produit de  $k(k-1)/2$  copules

bivariées à l'aide du conditionnement, une décomposition de type Vine. Dans le cas tri-varié, et comme dans tous les autres cas  $k$ -variés avec  $k \geq 3$ , la décomposition de type Vine n'est pas unique. En effet, on peut obtenir aussi dans le même cas tri-varié que

$$\begin{aligned} c(F_1(x_1), F_2(x_2), F_3(x_3)) &= c_{12}(F_1(x_1), F_2(x_2)) c_{13}(F_1(x_1), F_3(x_3)) \\ &\times c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1) | x_1). \end{aligned} \quad (2.7)$$

Ces deux décompositions correspondent respectivement à une décomposition D-Vine pour (2.6) et à une décomposition C-Vine pour (2.7) d'un cas trivarié. Le D-vine et le C-Vine font référence aux termes Dynamic Vine et Canonical Vine respectivement. Maintenant que l'idée derrière la décomposition de type Vine est bien exposée, le D-Vine et le C-Vine peuvent être décrites.

À partir d'ici, et pour le reste du mémoire, on simplifie les notations en posant  $f_i(x_i) = f_i$  et

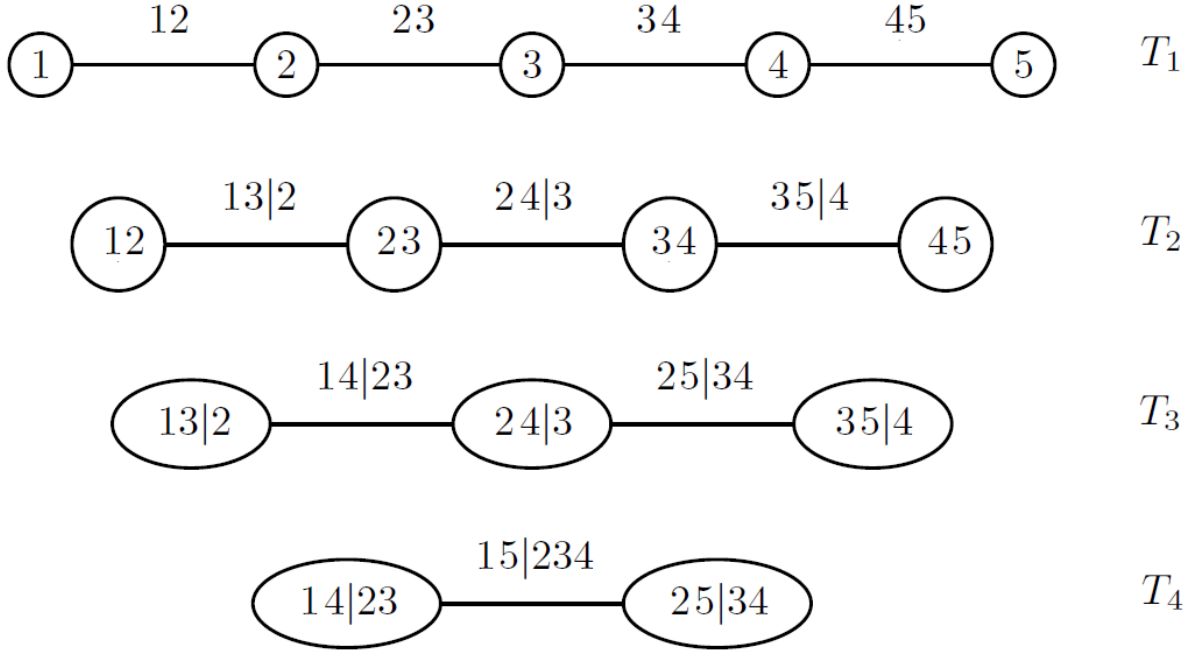
$$c_{ij|D}(F_{i|D}(x_i), F_{j|D}(x_j) | x_D) = c_{ij|D}.$$

Toutes les décompositions de type Vine pour une copule  $k$ -variée se représentent à partir de  $(k - 1)$  arbres qui, au final, décomposent le tout en  $k(k - 1)/2$  copules bivariées. Pour un D-Vine, comme celui à la figure 2.8 où  $k = 5$ , le premier arbre est construit en représentant chaque variable  $X_i$ ,  $i = 1, 2, \dots, k$  par un noeud. Chaque variable  $X_i$  est notée simplement  $i$  dans un noeud. Chaque noeud consécutif est relié par une branche telle que le noeud  $i$  est relié aux noeuds  $i - 1$  et  $i + 1$  pour  $i = 2, 3, \dots, (k - 1)$ . Chaque branche représente alors la structure de dépendance entre les noeuds qu'elle relie, c'est-à-dire une densité de copule. Ainsi, dans le premier arbre on note la  $r^e$  branche " $r(r + 1)$ " et elle représente la copule de densité  $c_{r(r+1)}$  avec  $r = 1, 2, \dots, (k - 1)$ . Chacune de ces branches, qui sont associées à des copules bivariées, font partie de la décomposition de type D-Vine : il s'agit des  $k - 1$  premières copules bivariées de la décomposition.

Pour chaque arbre qui suit, de l'arbre  $T_2$  à l'arbre  $T_{k-1}$ , on considère les branches de

l'arbre précédent comme les noeuds et on trace de nouvelles branches reliant les noeuds consécutifs. Pour tous les noeuds consécutifs  $ij|D$  et  $sl|D'$ , on considère  $D_1 = \{i, j, D\}$  et  $D_2 = \{s, l, D'\}$ , on pose  $D'' = D_1 \cap D_2$  et  $J = (D_1 \cup D_2) \setminus \{D''\} = \{j_1, j_2\}$  pour finalement obtenir que la branche reliant  $ij|D$  et  $sl|D'$  est donnée par  $j_1 j_2 | D''$  qui représente la copule bivariable pour les variables  $X_{j_1}$  et  $X_{j_2}$  conditionnelle aux variables  $X_{d''}$  avec  $d'' \in D''$ . Les  $k(k-1)/2$  branches des  $k-1$  arbres fournissent l'ensemble des copules bivariées constituant la décomposition D-Vine.

Figure 2.8 – Arbres de la décomposition D-Vine d'une copule  $k$ -variée avec  $k = 5$



Dans l'exemple de l'arbre à la figure 2.8, on constate que la décomposition D-Vine de la densité d'une copule  $k$ -variée avec  $k = 5$  notée  $c$  est donnée par

$$\begin{aligned}
c &= c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} && (\text{Arbre } T_1) \\
&\times c_{13|2} \cdot c_{24|3} \cdot c_{35|4} && (\text{Arbre } T_2) \\
&\times c_{14|23} \cdot c_{25|34} && (\text{Arbre } T_3) \\
&\times c_{15|234} && (\text{Arbre } T_4)
\end{aligned}$$

Prendre note que chaque copule bivariee possede un (ou parfois plusieurs) parametre(s) et qu'il a ete omis volontairement d'ecrire  $c_{ij|d;\phi}$  ou  $\phi$  est le parametre de la copule. Cette omission est due au fait que l'interet de cette section n'est pas d'estimer les parametres, mais plutot d'expliquer les decompositions D-Vine et C-Vine. De meme, la densite de copule  $k$ -variee  $c$  devrait s'ecrire  $c_\Omega$  ou  $\Omega$  est la reunion de tous les parametres des copules bivariees de la decomposition.

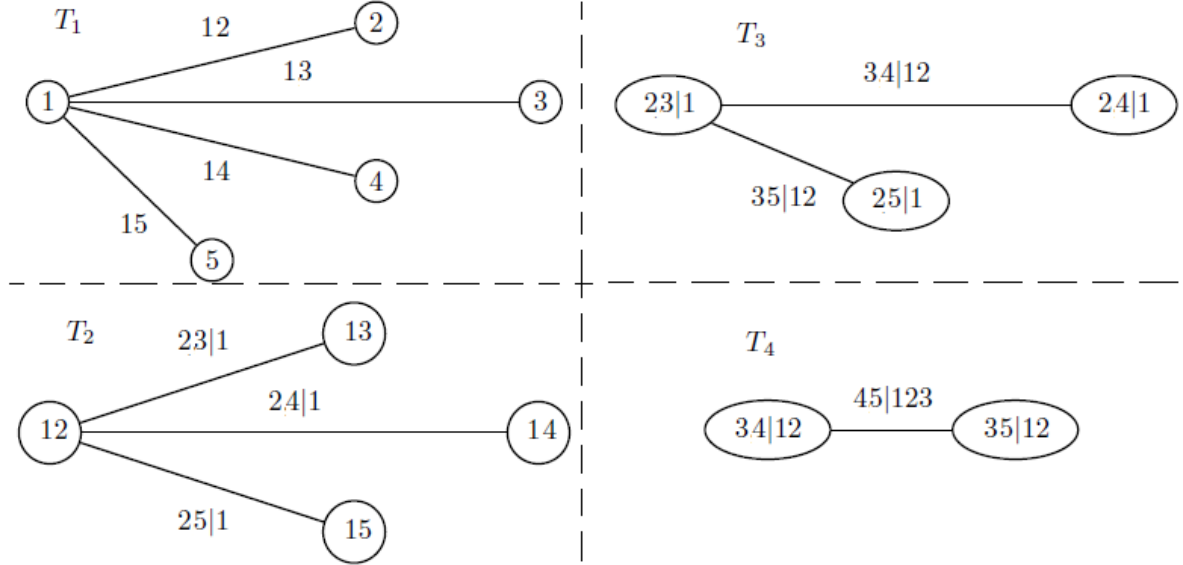
Pour la decomposition C-Vine, seule la construction des arbres change. Comme mentionne auparavant, toutes les decompositions de type Vine de copules  $k$ -variees comportent  $(k-1)$  arbres avec  $k(k-1)/2$  branches. Pour le C-Vine, on construit le premier arbre  $T_1$  en reliant le premier noeud, associe a la premiere variable aleatoire, notee "1", a chacun des autres noeuds. Ainsi, le noeud "1" est relie a tous les autres noeuds, du noeud "2" au noeud " $k$ ". Chacune de ces branches reliant le premier noeud aux autres noeuds correspond a une copule telle que la branche entre le noeuds 1 et  $i$  se nomme " $1i$ " en reference a la copule  $c_{1i}$ . Pour la construction de l'arbre  $T_h$  avec  $h$  allant de 2 a  $(k-1)$ , la procedure est la suivante :

1. Toutes les branches de l'arbre  $T_{h-1}$  correspondent a un noeud de l'arbre  $T_h$  et la premiere branche de l'arbre  $T_{h-1}$  devient le premier noeud de l'arbre  $T_h$ .
2. On relie le premier noeud de  $T_h$  a tous les autres noeuds pour former  $(k-h)$  branches chacune etant associee a une copule bivariee.
3. On determine chaque branche de la meme facon qu'avec le D-Vine, c'est-a-dire que la branche entre  $ij|D$  et  $sl|D'$  est donnee par  $j_1j_2|D''$  ou  $D'' = D_1 \cap D_2$  et  $J = (D_1 \cup D_2) \setminus \{D''\} = \{j_1, j_2\}$  ou  $D_1 = \{i, j, D\}$  et  $D_2 = \{s, l, D'\}$ .

Un exemple de construction d'arbres pour la decomposition C-Vine d'une copule  $k$ -variee avec  $k = 5$  est donne a la figure 2.9. Dans cet exemple, on obtient la decomposition C-Vine suivante pour la copule multivariee  $c$  :

$$\begin{aligned}
c &= c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{15} & (\text{Arbre } T_1) \\
\times & c_{23|1} \cdot c_{24|1} \cdot c_{25|1} & (\text{Arbre } T_2) \\
\times & c_{34|12} \cdot c_{35|12} & (\text{Arbre } T_3) \\
\times & c_{45|123} & (\text{Arbre } T_4)
\end{aligned}$$

Figure 2.9 – Arbres de la décomposition C-Vine d'une copule  $k$ -variée avec  $k = 5$



Peu importe la forme de la décomposition de type Vine, de façon générale la décomposition d'une copule  $k$ -variée  $c$  s'écrit

$$c = \prod_{h=1}^{k-1} \prod_{e \in E_h} c_{a_e, b_e | D_e}, \quad (2.8)$$

où  $E_h$  est l'ensemble des branches de l'arbre  $T_h$ , et  $a_e$ ,  $b_e$  et  $D_e$  sont respectivement le premier terme, le deuxième terme et l'ensemble des indices de conditionnement pour la branche  $e \in E_h$  avec  $h = 1, 2, \dots, (k-1)$ . On note que la cardinalité de l'ensemble  $E_h$  est donnée par  $(k-h)$ .

Le D-Vine et le C-Vine sont les décompositions les plus communes dans la littérature mathématique, mais d'autres décompositions existent comme le M-Vine présenté à la section suivante.

### 2.4.2 M-Vine

La décomposition M-Vine a été introduite pour les cas spécifiques aux séries chronologiques. Soit  $\mathbf{X}_t = (X_{1;t}, \dots, X_{k;t})$  une série chronologique  $k$ -variée et markovienne d'ordre  $m$  fortement stationnaire. On note " $ij$ " la variable aléatoire associée à  $X_{i;t+j-m}$  avec  $i = 1, \dots, k$  et  $j = 1, \dots, m$ . La construction des arbres de la décomposition M-Vine peut maintenant être définie. Pour le premier arbre, les noeuds notés par " $ij$ " représente la densité de la variable  $X_{i;t+j-m}$ . Quant aux branches, il y en a entre " $1j$ " et " $1(j+1)$ " pour chaque  $j = 1, 2, \dots, (m-1)$  et entre " $ij$ " et " $(i+1)j$ " pour chaque  $i = 1, 2, \dots, (k-1)$  et  $j = 1, \dots, m$ . Chaque branche de ce premier arbre est identifiée de la même façon que pour le C-Vine et le D-Vine, c'est-à-dire que la branche entre les noeuds " $ij$ " et " $kl$ " est donnée par " $ij, kl$ " qui correspond alors à la copule  $c_{ij,kl}$  représentant la structure de dépendance entre les variables  $X_{i;t+j-m}$  et  $X_{k;t+l-m}$ .

Pour le deuxième arbre, toutes les branches du premier arbre deviennent les noeuds du deuxième. Chaque noeud " $ij, sl$ " de l'arbre  $T_2$  est lié aux noeuds " $i'j', s'l'$ " pour lesquels les branches associées dans l'arbre précédent  $T_1$  avaient un noeud commun. Par exemple, dans la figure 2.10, le noeud " $11,21$ " de l'arbre  $T_2$  est lié par des branches aux noeuds " $21,31$ " et " $11,12$ ". Cela est dû au fait que la branche " $11,21$ " dans l'arbre  $T_1$  relie les noeuds " $11$ " et " $21$ " touchés respectivement par les branches " $11,12$ " et " $21,31$ ". Il y a une exception à cette règle : on ne trace pas de branche entre les noeuds " $1(j-1), 1j$ " et " $1j, 1(j+1)$ " avec  $j = 2, \dots, (m-1)$  pour l'arbre  $T_2$ . En effet, même si les branches analogues de ces noeuds de  $T_2$  touchent toutes les deux au noeud " $1j$ " dans  $T_1$ , on y trace pas de branches. Cette dernière remarque est valide pour tous les duos de noeuds " $1(j-1), 1j$ " et " $1j, 1(j+1)$ " dans  $T_2$  pour  $j = 2, \dots, (m-1)$ . On remarque cette exception dans l'arbre  $T_2$  de la figure 2.10 dans lequel il n'y a pas de branche entre les noeuds " $11,12$ " et " $12,13$ ". Cela est dû au fait que les noeuds " $11,12$ " et " $12,13$ " de  $T_2$  sont de la forme " $1(j-1), 1j$ " et " $1j, 1(j+1)$ " avec  $j = 2$ . Chaque branche tracée pour

l'arbre  $T_2$  est décrite comme dans le cas des décompositions C-Vine et D-Vine.

Pour la construction de l'arbre  $T_h$ , de la même façon que pour l'arbre  $T_2$ , toutes les branches de  $T_{h-1}$  deviennent les noeuds de  $T_h$ . Ensuite, on relie tous les noeuds de  $T_h$  pour lesquels les branches de  $T_{h-1}$  associées ont un noeud commun. Tout comme pour la construction de  $T_2$ , il y a une exception à cette dernière règle : quand un noeud de  $T_{h-1}$  est atteint par 3 branches, seul 2 des 3 noeuds associés dans  $T_h$  à ces 3 branches de  $T_{h-1}$  sont reliés par une branche. Ce sont les 2 noeuds " $ij, sl|D$ " et " $i'j', s'l'|D'$ " avec  $D = D'$  qui ne sont pas liés. Un exemple est fourni par l'arbre  $T_3$  de la figure 2.10 dans lequel les noeuds " $11,22|12$ " et " $13,22|12$ " ne sont pas liés d'une branche alors qu'on y remarque que leurs branches associées dans  $T_2$  relié par le noeud " $12,22$ ", mais que ce dernier noeud possède 3 branches. Parmi les trois branches, on constate que les branches " $11,22|12$ " et " $13,22|12$ " sont de la forme " $ij, sl|D$ " et " $i'j', s'l'|D'$ " avec  $D = D'$  où  $D = D' = 12$ . Donc, dans l'arbre  $T_3$ , on ne doit pas tracer de branche entre les noeuds " $11,22|12$ " et " $13,22|12$ ". Enfin, chaque branche tracée pour l'arbre  $T_h$  est décrite de la même façon que dans le cas des décompositions C-Vine et D-Vine.

Les  $k(m+1)(k(m+1)-1)/2$  branches, associées à des copules bivariées, des  $(k(m+1)-1)$  arbres fournissent la décomposition M-Vine de la copule  $c$  modélisant la dépendance de  $\mathbf{X}_t$ . L'expression (2.8) est vérifié et le produit des copules bivariées est la copule  $c$ .

La figure 2.10 procure un exemple des arbres  $T_1$ ,  $T_2$ ,  $T_3$ ,  $T_7$  et  $T_8$  obtenus lors d'une décomposition M-Vine d'une série chronologique markovienne d'ordre 2 trivariée

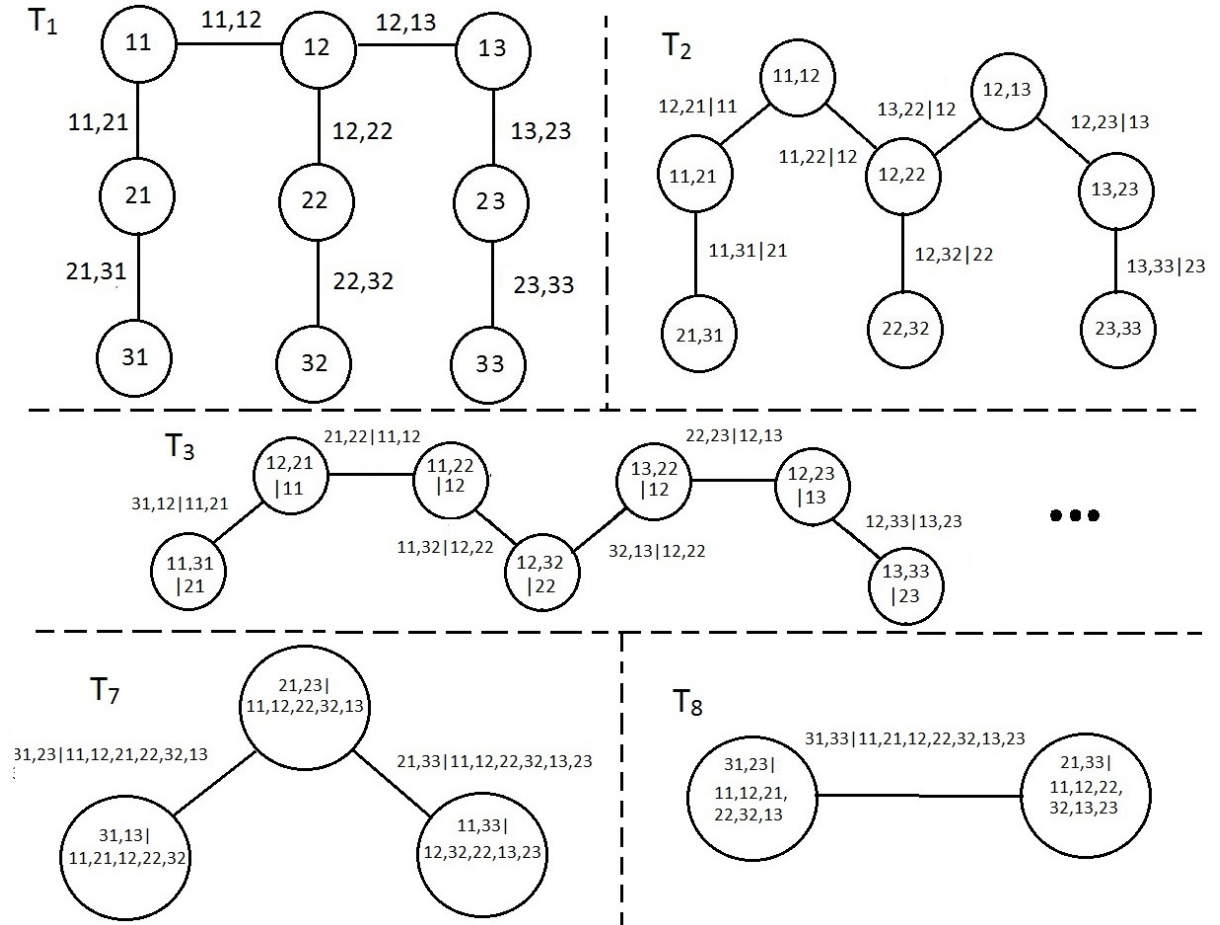
$$\mathbf{X}_t = (X_{1;t-2}, X_{2;t-2}, X_{3;t-2}, X_{1;t-1}, X_{2;t-1}, X_{3;t-1}, X_{1;t}, X_{2;t}, X_{3;t}).$$

On en déduit que la décomposition M-Vine de la copule  $c$  représentant la structure de dépendance de cette série chronologique est donné par :



$$\begin{aligned}
c &= c_{21,31} \cdot c_{11,21} \cdot c_{11,12} \cdot c_{12,22} \cdot c_{22,32} \cdot c_{12,13} \cdot c_{13,23} \cdot c_{23,33} & (\text{Arbre } T_1) \\
&\times c_{11,31|21} \cdot c_{12,21|11} \cdot c_{11,22|12} \cdot c_{12,32|22} \cdot c_{13,22|12} \cdot c_{12,23|13} \cdot c_{13,33|23} & (\text{Arbre } T_2) \\
&\times c_{31,12|11,21} \cdot c_{21,22|11,12} \cdot c_{11,32|12,22} \cdot c_{32,13|12,22} \cdot c_{22,23|12,13} \cdot c_{12,33|13,23} & (\text{Arbre } T_3) \\
&\times c_{31,22|11,21,12} \cdot c_{21,32|11,12,22} \cdot c_{11,13|12,22,32} \cdot c_{32,23|12,13,22} \cdot c_{22,33|12,13,23} & (\text{Arbre } T_4) \\
&\times c_{31,32|11,21,12,22} \cdot c_{21,13|11,12,22,32} \cdot c_{11,23|12,13,22,32} \cdot c_{32,33|12,22,13,23} & (\text{Arbre } T_5) \\
&\times c_{31,13|11,21,12,22,32} \cdot c_{21,23|11,12,22,32,13} \cdot c_{11,33|12,22,32,13,23} & (\text{Arbre } T_6) \\
&\times c_{31,23|11,21,12,22,32,13} \cdot c_{21,33|11,12,22,32,13,23} & (\text{Arbre } T_7) \\
&\times c_{31,33|11,21,12,22,32,13,23} & (\text{Arbre } T_8)
\end{aligned}$$

Figure 2.10 – Les arbres  $T_1$ ,  $T_2$ ,  $T_3$ ,  $T_7$  et  $T_8$  de la décomposition M-Vine d'une copule associée à une série chronologique markovienne d'ordre 3 tri-variee



# CHAPITRE 3

## Lien entre VAR(1) et M-Vine

Le processus VAR(1) et la décomposition M-Vine (M-V) abordés dans les précédentes sections se veulent deux manières de représenter des séries chronologiques multivariées stationnaires et la dépendance intra et inter séries. Cependant, le modèle M-V se veut un moyen pour représenter la structure de dépendance plus général que le modèle VAR(1) stationnaire, dit classique. Comme on s'intéresse à une équivalence entre VAR(1) et M-V et que le M-V ne peut être utilisé quand dans le cas stationnaire, on ne considérera que le modèle VAR(1) stationnaire. Bref, il devrait être possible de trouver la représentation M-V d'un processus VAR(1) puisque le processus VAR(1) devient un cas particulier du M-V. L'objectif de cette présente section sera justement de définir les liens entre ces deux représentations en montrant comment passer d'une représentation à une autre.

Le cas univarié AR(1) de ce problème a déjà été abordé par Chen et Fan (2006) [CF06]. En effet, ils ont spécifié que si une série chronologique  $Y_t$  stationnaire avec une distribution marginale  $G^*(\cdot)$  et une structure de dépendance définie à l'aide une copule gaussienne bivariable  $C(\cdot, \cdot; \alpha)$  avec un paramètre  $\alpha \in (-1, 1)$  :

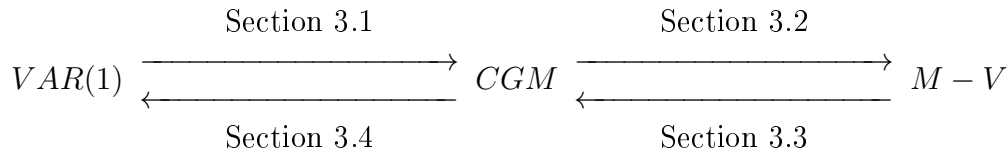
$$C(v_1, v_2; \alpha) = \varphi_\alpha(\varphi^{-1}(v_1), \varphi^{-1}(v_2)),$$

où  $\varphi$  est la fonction de répartition d'une loi normale centrée réduite et  $\varphi_\alpha$  est la fonction de répartition d'une loi normale bivariée avec paramètre de corrélation  $\alpha$ , alors  $Z_t = \{\varphi^{-1}(G^*(Y_t))\}$  est un processus gaussien représenté par  $Z_t = \alpha Z_{t-1} + \epsilon_t$  et  $\epsilon_t$  est indépendant de  $Y_{t-1}$ . Si la distribution marginale est une normale centrée réduite, un modèle autorégressif s'applique à  $Y_t$ ; sinon  $Z_t$  est un processus markovien du premier ordre.

En résumé, Chen et Fan (2006) [CF06] ont montré qu'une série chronologique stationnaire dont la structure de dépendance est caractérisée par une copule peut être représentée à l'aide d'un modèle AR(1) sous certaines conditions et inversement.

Les sections du présent chapitre aborde de façon détaillée le cas multidimensionnel, c'est-à-dire la décomposition M-V d'un VAR(1) et la reconstruction d'un VAR(1) à partir de sa décomposition M-V. Les passages de VAR(1) à M-V et de M-V à VAR(1) se font chacun en passant d'abord par une représentation avec la copule gaussienne multivariée (CGM). On rappelle au lecteur qu'un manuscrit résumant les contributions de ce chapitre a été soumis pour publication à la revue Statistics and Probability Letters et est reproduit en annexe. Les sous-sections décrivant les différentes transitions entre les représentations VAR(1), CGM et M-V sont présentées selon ce schéma :

Figure 3.1 – Schéma des transitions de représentation avec leur section associée



### 3.1 Représentation de la structure de dépendance d'un VAR(1) avec une copule gaussienne multivariée

Soit  $\mathbf{X}_t = (X_{1;t}, X_{2;t}, \dots, X_{k;t})^T$  une série chronologique stationnaire VAR(1)  $k$ -variée définie par

$$\mathbf{X}_t = \mathbf{v} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t, \quad \text{avec } t \in \mathbb{N}, \quad (3.1)$$

où  $\mathbf{v}$  est un vecteur de longueur  $k$  de constantes,  $\Phi$  est une matrice de coefficients de format  $k \times k$  et  $\boldsymbol{\epsilon}_t, t \in \mathbb{N}$ , sont des vecteurs  $k$ -variés i.i.d. normalement distribués avec moyenne zéro et matrice de variance-covariance  $\Sigma$ .

Pour trouver une représentation CGM pour ce VAR(1), on calcule la distribution conjointe de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$ . On sait que la distribution de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$  est gaussienne puisque, avec la stationnarité et le fait que

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{v} + \Phi \mathbf{X}_0 + \boldsymbol{\epsilon}_1 \\ \mathbf{X}_2 &= \mathbf{v} + \Phi \mathbf{X}_1 + \boldsymbol{\epsilon}_2 \\ &\vdots \\ \mathbf{X}_{t-1} &= \mathbf{v} + \Phi \mathbf{X}_{t-2} + \boldsymbol{\epsilon}_{t-1} \\ \mathbf{X}_t &= \mathbf{v} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t, \end{aligned}$$

on a que chaque réalisation  $\mathbf{X}_t$  est une somme de la réalisation précédente  $\mathbf{X}_{t-1}$  et d'une innovation gaussienne  $\boldsymbol{\epsilon}_t$ . Cette réalisation précédente est elle-même la somme de sa propre réalisation précédente et d'une innovation gaussienne. Comme les innovations  $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_t$  sont i.i.d.  $N(0, \Sigma)$ , on a que toutes les réalisations  $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_{t-1}, \mathbf{X}_t$  sont gaussiennes.

Aussi, comme on a que  $(\mathbf{X}_t | \mathbf{X}_{t-1}) \sim N(\mathbf{v} + \Phi \mathbf{X}_{t-1}, \Sigma)$ , que  $\mathbf{X}_{t-1}$  suit une loi gaussienne, que  $(\mathbf{X}_t | \mathbf{X}_{t-1})$  et  $\mathbf{X}_{t-1}$  sont indépendantes et que  $f(x_t, x_{t-1}) = h(x_t | x_{t-1}) g(x_{t-1})$  où  $f$ ,

$h$  et  $g$  sont la densité conjointe de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$ , la densité conditionnelle de  $\mathbf{X}_t|\mathbf{X}_{t-1}$  et la densité marginale de  $\mathbf{X}_{t-1}$ , on peut affirmer que la loi conjointe de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$  est gaussienne.

Aussi, comme  $\mathbf{X}_t$  est stationnaire, on a que :

1. Les racines de  $\det(\mathbb{I} - \Phi) = 0$  sont supérieures à 1 et donc que  $(\mathbb{I} - \Phi)$  est inversible ;
2.  $E[\mathbf{X}_t] = E[\mathbf{X}_{t-j}], \forall j \in \mathbb{Z}$  ;
3.  $Var(\mathbf{X}_t) = Var(\mathbf{X}_{t-j}), \forall j \in \mathbb{Z}$  ;
4.  $Cov(\mathbf{X}_t, \mathbf{X}_{t-j}) = Cov(\mathbf{X}_{t'}, \mathbf{X}_{t'-j}), \forall i \in \mathbb{Z}$ .

On déduit de ces dernières observations que :

$$\begin{aligned}
 E[\mathbf{X}_{t-1}] &= E[\mathbf{X}_t] \\
 &= E[\mathbf{v} + \Phi\mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t] \\
 &= \mathbf{v} + \Phi E[\mathbf{X}_{t-1}] \\
 &= (\mathbb{I} - \Phi)^{-1} \mathbf{v}.
 \end{aligned}$$

Ensuite, en posant  $A = Var(X_t) = Var(X_{t-1})$  on note que :

$$\begin{aligned}
 Cov(\mathbf{X}_t, \mathbf{X}_{t-1}) &= Cov(\mathbf{v} + \Phi\mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t, \mathbf{X}_{t-1}) \\
 &= \Phi Var(\mathbf{X}_{t-1}) + Cov(\boldsymbol{\epsilon}_t, \mathbf{X}_{t-1}) \\
 &= \Phi A.
 \end{aligned}$$

Bref, on obtient que :

$$\begin{pmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_t \end{pmatrix} \sim N_{2k} \left( \begin{pmatrix} (\mathbb{I} - \Phi)^{-1} \mathbf{v} \\ (\mathbb{I} - \Phi)^{-1} \mathbf{v} \end{pmatrix}, \begin{pmatrix} A & \Phi A \\ A\Phi^T & A \end{pmatrix} \right)$$

Comme  $\mathbf{X}_t$  est stationnaire, on a l'équivalence suivante :

$$\begin{aligned} Var(\mathbf{X}_{t-1}) &= Var(\mathbf{X}_t) \\ &= Var(\mathbf{v} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t) \\ &= \Phi Var(\mathbf{X}_{t-1}) \Phi^T + \Sigma. \end{aligned}$$

Donc, avec  $A = Var(\mathbf{X}_t)$ , la stationnarité de  $\mathbf{X}_t$  est vérifiée si et seulement si  $\Sigma$  est tel que

$$\Sigma = A - \Phi A \Phi^T$$

Pour trouver une expression explicite à  $A = Var(\mathbf{X}_t)$ , on considère l'opérateur  $vec(\cdot)$  de vectorisation de matrice tel que

$$vec\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

L'une des propriétés de l'opérateur  $vec(\cdot)$  est que

$$vec(ABC) = (C^T \otimes A) vec(B)$$

où  $\otimes$  est le produit de matrices de Kronecker. Pour davantage d'informations sur l'opérateur  $vec(\cdot)$  et sur le produit de Kronecker, on réfère le lecteur à [MO13] et à [HS81] respectivement.

On rappelle que le produit de matrices de Kronecker est défini tel que pour des matrices  $A$  de format  $m \times p$  et  $B$  de format  $n \times q$  on a :

$$C = A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1p}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mp}B \end{pmatrix}$$

où  $C$  est alors de format  $mn \times pq$ .

Donc, comme  $A = \Phi A \Phi^T + \Sigma$ , on a que :

$$\begin{aligned} \text{vec}(A) &= \text{vec}(\Phi A \Phi^T + \Sigma) \\ &= \text{vec}(\Phi A \Phi^T) + \text{vec}(\Sigma) \\ &= (\Phi \otimes \Phi) \text{vec}(A) + \text{vec}(\Sigma) \\ &= (\mathbb{I} - D)^{-1} \text{vec}(\Sigma), \end{aligned}$$

où  $D = \Phi \otimes \Phi$  est de format  $k^2 \times k^2$ .

L'utilisation de l'opérateur  $\text{vec}(\cdot)$  et du produit de matrices de Kronecker pour résoudre ce système est basé sur l'approche utilisée par Luca Gambetti dans plusieurs de ses travaux dont son article [FGS19].

On peut ensuite déduire du vecteur  $k^2$ -varié  $\text{vec}(A)$  la matrice  $A$  de format  $k \times k$ .

Finalement, on a que la loi de conjointe de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$  est une normale  $2k$ -varié de moyenne  $\boldsymbol{\mu}$  et de matrice de variance covariance  $\bar{\Sigma}$  :

$$\boldsymbol{\mu} = \begin{pmatrix} (\mathbb{I} - \Phi)^{-1} \mathbf{v} \\ (\mathbb{I} - \Phi)^{-1} \mathbf{v} \end{pmatrix} \quad \text{et} \quad \bar{\Sigma} = \begin{pmatrix} A & \Phi A \\ A \Phi^T & A \end{pmatrix}.$$

À partir de la densité conjointe de  $\mathbf{X}_t$  et  $\mathbf{X}_{t-1}$ , on a que toutes les densités marginales sont gaussiennes, c'est-à-dire qu'on a  $X_{i;t} \sim N(\bar{\mu}_i, \bar{\Sigma}_{i,i})$  pour  $i = 1, \dots, k$ . De plus, la copule multivariée qui décrit la structure de dépendance de  $(\mathbf{X}_t^T, \mathbf{X}_{t-1}^T)^T$  est une copule gaussienne. Cette CGM possède  $k(k-1)/2$  paramètres,  $\rho_{ih,jl}$ , qui représente la corrélation entre la  $i^e$  composante de  $\mathbf{X}_{t+h-2}$  et la  $j^e$  composante de  $\mathbf{X}_{t+l-2}$  avec  $i, j \in \{1, \dots, k\}$  et  $h, l \in \{1, 2\}$ ; prendre note que pour un modèle VAR(1), qui est markovien d'ordre 1,  $\rho_{i1,j1}$  est égal à  $\rho_{i2,j2}$ . Pour obtenir la CGM, tous les paramètres de corrélation  $\rho_{ih,jl}$  peuvent être calculés à partir de  $\bar{\Sigma}$  de cette façon :

$$\rho_{ih,jl} = \frac{\bar{\Sigma}_{(2-h)k+i, (2-l)k+j}}{\bar{\Sigma}_{i,i} \bar{\Sigma}_{j,j}}. \quad (3.2)$$

## 3.2 Décomposition M-Vine d'une copule gaussienne multivariée

Un objectif important de ce mémoire est d'étudier en détail la décomposition du processus VAR(1) en M-V. Pour ce faire, et parce qu'il a été montré dans la Section 3.1 qu'un modèle VAR(1) peut se représenter comme un CGM sous certaines conditions, on décrit le passage d'une représentation CGM à une représentation M-Vine pour une structure de dépendance en vue d'obtenir une décomposition M-Vine du processus VAR(1).

Dans la Section 3.1, on a montré que la structure de dépendance d'un processus VAR(1)  $k$ -varié,  $\mathbf{X}_t$ ,  $t \in \mathbb{N}$ , peut être caractérisée par une  $2k$ -dimensionnelle CGM, notée  $c_\Omega$ , où  $\Omega$  est l'ensemble des  $d = k(2k - 1)$  paramètres. Le M-V décompose une  $2k$ -dimensionnelle CGM en  $d$  copules bivariées avec  $2k - 1$  copules non-conditionnelles dans le premier arbre. Prendre note, qu'une copule non-conditionnelle sera dite directe. Toutes les autres copules, du deuxième au  $(2k - 1)^e$  arbre, sont conditionnelles. Chaque copule bivariée correspond à une branche d'un arbre dans la décomposition M-V.

En considérant seulement les modèles CGM qui sont caractérisés par  $d$  paramètres provenant de l'ensemble  $\Omega$ , on a que les paramètres  $\rho_{ih,jl}$  représente les corrélations entre  $X_{i;t+h-2}$  et  $X_{j;t+l-2}$  où  $i, j \in \{1, \dots, k\}$  et  $h, l \in \{1, 2\}$ . Avec une CGM, les copules bivariées directes et conditionnelles de sa décomposition M-V n'auront pas d'autre choix que d'être gaussienne. En effet, on rappelle qu'avec  $E_w$ , l'ensemble des branches du  $w^e$  arbre, où  $e = c_{a_e, b_e | D_e} \in E_w$  est une branche telle que définie à la Section 2.4.2, on a que

$$c_\Omega = \prod_{w=1}^{2k-1} \prod_{e \in E_w} c_{a_e, b_e | D_e} \quad (3.3)$$

et que seul un produit de copules gaussiennes  $(c_{a_e, b_e | D_e})$  peut donner une CGM  $(c_\Omega)$ .



Cela implique que toutes les copules bivariées de la décomposition M-V d'une CGM sont gaussiennes. Donc, tous les paramètres du modèle M-V sont des corrélations directes et conditionnelles. On note que toutes les corrélations directes du M-V, c'est-à-dire les paramètres du premier arbre, sont directement données par la CGM. Ainsi, pour obtenir les paramètres pour le M-V, les corrélations conditionnelles associées à chaque branche de chaque arbre, du deuxième au  $(2k - 1)^e$ , sont requises. Pour les obtenir, la théorie sur les distributions gaussiennes conditionnelles sera utilisée.

Tout d'abord, on définit un opérateur pour les matrices carrées, que l'on note  $T'$ , qui simplifiera les prochains développements. Soit  $M = [m_{i;j}]_{i,j=1,\dots,k}$  une matrice carrée de format  $k \times k$ . Alors,  $M^{T'} = [m'_{i;j}]_{i,j=1,\dots,k}$  où  $m'_{i;j} = m_{k+1-j;k+1-i}$  pour  $i, j = 1, \dots, k$ . Par exemple, on a

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{T'} = \begin{bmatrix} i & f & c \\ h & e & b \\ g & d & a \end{bmatrix}.$$

Pour optimiser les calculs récurrents utilisés afin d'obtenir les paramètres pour le M-V, c'est-à-dire des corrélations conditionnelles, on considère une matrice  $\Psi$  qui contient tous les paramètres de la CGM, c'est-à-dire les corrélations directes, organisés comme suit :

$$\Psi = \begin{pmatrix} U & V \\ V^T & U^{T'} \end{pmatrix} \quad (3.4)$$

où

$$U = \begin{pmatrix} 1 & \rho_{(k-1)1,k1} & \cdots & \rho_{11,k1} \\ \rho_{(k-1)1,k1} & 1 & \cdots & \rho_{11,(k-1)1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{11,k1} & \rho_{11,(k-1)1} & \cdots & 1 \end{pmatrix} \text{ et } V = \begin{pmatrix} \rho_{k1,12} & \cdots & \rho_{k1,k2} \\ \rho_{(k-1)1,12} & \cdots & \rho_{(k-1)1,k2} \\ \vdots & & \vdots \\ \rho_{11,12} & \cdots & \rho_{11,k2} \end{pmatrix}.$$

La théorie sur les distributions gaussiennes conditionnelles indique qu'avec

$$Z = (\mathbf{X}^T, \mathbf{Y}^T)^T \sim N(\boldsymbol{\mu}, \Sigma)$$

où  $\boldsymbol{\mu} = (\boldsymbol{\mu}_{\mathbf{X}}^T, \boldsymbol{\mu}_{\mathbf{Y}}^T)^T$  et  $\Sigma = \begin{pmatrix} \Sigma_{\mathbf{X}} & \Sigma_{\mathbf{X},\mathbf{Y}} \\ \Sigma_{\mathbf{Y},\mathbf{X}} & \Sigma_{\mathbf{Y}} \end{pmatrix}$ , on a :

$$\begin{aligned} \boldsymbol{\mu}_{\mathbf{X}|\mathbf{Y}=\mathbf{a}} &= \boldsymbol{\mu}_{\mathbf{X}} + \Sigma_{\mathbf{X},\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}(\mathbf{a} - \boldsymbol{\mu}_{\mathbf{Y}}) \\ \Sigma_{\mathbf{X}|\mathbf{Y}=\mathbf{a}} &= \Sigma_{\mathbf{X}} - \Sigma_{\mathbf{X},\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}\Sigma_{\mathbf{Y},\mathbf{X}}. \end{aligned} \quad (3.5)$$

Comme les corrélations sont invariantes aux changements d'échelle et d'origine, on assume, pour le reste de ce chapitre, que les distributions marginales sont des gaussiennes centrées réduites quand on fait les calculs de corrélations conditionnelles (et de corrélations directes dans la Section 3.3).

Pour calculer les entrées de la matrice de corrélation conditionnelle, on utilise l'équation (3.5) d'un façon récursive :

$$\psi^{j+1,b} = \text{diag}(\boldsymbol{\Psi}^{j,b})^{-1/2} (\boldsymbol{\Psi}^{j,b}) \text{diag}(\boldsymbol{\Psi}^{j,b})^{-1/2}, \quad (3.6)$$

où, pour  $b = 2, \dots, 2k - 1$  et  $j = 0, \dots, 2k - b$ ,  $\boldsymbol{\Psi}^{j,b} = \psi_1^{j,b} - \psi_2^{j,b} \left( \psi_2^{j,b} \right)^T$ ,  $\psi_1^{j,b}$  est égal à la matrice  $\psi^{j,b}$  sans sa  $b^e$  ligne et sa  $b^e$  colonne et  $\psi_2^{j,b}$  est la  $b^e$  colonne de  $\psi^{j,b}$  sans son  $b^e$  terme (c'est-à-dire l'entrée égale à 1 dans la colonne) avec  $\psi^{0,b} = \boldsymbol{\Psi}$ ; et  $\text{diag}(\cdot)$  est un opérateur pour les matrices carrées qui retourne une matrice de même format avec seulement la diagonale non-nulle.

L'équation (3.6) est très pratique et donne toutes les corrélations conditionnelles nécessaires pour la décomposition M-V d'une CGM  $2k$ -dimensionnelle. Tous les paramètres du M-V peuvent être obtenus à partir d'une CGM en deux étapes :

étape 1 : La matrice de corrélations initiale contenant les paramètres de la CGM,  $\boldsymbol{\Psi}$ , donne les corrélations directes de la décomposition M-Vine.

étape 2 : Pour  $b$  allant de 2 jusqu'à  $2k - 1$ ,

1.  $\psi^{0,b}$  est posée égale à  $\Psi$ .
2. Pour  $j$  allant de 1 jusqu'à  $2k - b$ ,
  - (a) Calculer  $\psi^{j,b}$  en utilisant l'équation (3.6) ;
  - (b)  $\psi_{(b-1),b}^{j,b}$  est un paramètre du M-V.

Éventuellement, tous les paramètres sont obtenus dans cet ordre :  $\rho_{(k-1)1,k1}$ ,  $\rho_{(k-2)1,(k-1)1}$ , ...,  $\rho_{21,11}$ ,  $\rho_{11,12}$ ,  $\rho_{12,22}$ , ...,  $\rho_{(k-1)2,k2}$ ,  $\rho_{(k-2)1,k1|(k-1)1}$ , ...,  $\rho_{11,k1|(k-1)1,(k-2)1,\dots,21}$ ,  $\rho_{k1,12|(k-1)1,\dots,11}$ , ...,  $\rho_{k1,k2|(k-1)1,\dots,11,12,\dots,(k-1)2}$ ,  $\rho_{(k-3)1,(k-1)1|(k-2)1,\dots,\rho_{(k-2)2,k2|(k-1)2}$ . Alors, tous les paramètres des copules gaussiennes bivariées impliquées dans la décomposition M-V peuvent être déduits du modèle CGM. Donc, un modèle VAR(1) peut être représenté avec un M-V puisque dans la Section 3.1, on a montré qu'un modèle VAR(1) peut être réécrit comme une CGM.

Pour bien illustrer le dernier calcul récursif, prenons l'exemple d'une série chronologique  $\mathbf{X}_t = (X_{1,t}, X_{2,t})$  bivariée markovienne d'ordre 1 dont la structure de dépendance est donnée par une CGM. Ainsi, on connaît tous les paramètres de la CGM, c'est-à-dire toutes les corrélations directes, qu'on regroupe dans la matrice  $\Psi$  comme suit :

$$\Psi = \begin{pmatrix} 1 & \rho_{21,11} & \rho_{21,12} & \rho_{21,22} \\ \rho_{21,11} & 1 & \rho_{11,12} & \rho_{11,22} \\ \rho_{21,12} & \rho_{11,12} & 1 & \rho_{11,21} \\ \rho_{21,22} & \rho_{11,22} & \rho_{11,21} & 1 \end{pmatrix}$$

Donc, on veut trouver tous les paramètres de la décomposition M-Vine pour  $\mathbf{X}_t$ , c'est-à-dire  $\rho_{21,11}$ ,  $\rho_{11,12}$ ,  $\rho_{21,12|11}$ ,  $\rho_{11,22|12}$  et  $\rho_{21,22|11,12}$ . La représentation CGM donne déjà les corrélations directes. Il ne reste qu'à trouver les corrélations conditionnelles. Pour trouver  $\rho_{21,12|11}$ , considérons  $\Psi$  comme la matrice de variance covariance de  $(\mathbf{X}_{t-1}, \mathbf{X}_t)$  et trouvons la matrice de variance covariance conditionnelle de  $(X_{2;t-1}, X_{1;t}, X_{2;t}|X_{1;t-1})$ .

On sépare  $\Psi$  ainsi :

$$\Psi = \left( \begin{array}{c|c|c|c} 1 & \rho_{21,11} & \rho_{21,12} & \rho_{21,22} \\ \hline \rho_{21,11} & 1 & \rho_{11,12} & \rho_{11,22} \\ \hline \rho_{21,12} & \rho_{11,12} & 1 & \rho_{11,21} \\ \hline \rho_{21,22} & \rho_{11,22} & \rho_{11,21} & 1 \end{array} \right)$$

Le but des prochaines manoeuvres est de trouver la matrice  $\Psi^{j=1,b=2}$ , où  $b$  signifie qu'on conditionne sur la  $b^e$  variable et  $j$  signifie que c'est la  $j^e$  fois qu'on conditionne à partir de la  $b^e$  variable, qui donne la matrice de variance covariance conditionnelle voulue. À partir de  $\Psi^{1,2}$ , il sera ensuite possible de trouver  $\psi^{1,2}$ , la matrice de corrélation conditionnelle. On note que  $\Psi^{0,b} = \psi^{0,b} = \Psi \ \forall b \in \{2,3\}$ . Tout d'abord, on pose :

$$\begin{aligned} \psi_1^{0,2} &= \begin{pmatrix} 1 & \rho_{21,12} & \rho_{21,22} \\ \rho_{21,12} & 1 & \rho_{11,21} \\ \rho_{21,22} & \rho_{11,21} & 1 \end{pmatrix}, \\ (\psi_2^{0,2})^T &= \begin{pmatrix} \rho_{21,11} & \rho_{11,12} & \rho_{11,22} \end{pmatrix}, \end{aligned}$$

à partir desquelles on obtient :

$$\begin{aligned} \Psi^{1,2} &= \psi_1^{0,2} - \psi_2^{0,2} (\psi_2^{0,2})^T \\ &= \begin{pmatrix} 1 - \rho_{21,11}^2 & \rho_{21,12} - \rho_{21,11}\rho_{11,12} & \rho_{21,22} - \rho_{21,11}\rho_{11,22} \\ \rho_{21,12} - \rho_{21,11}\rho_{11,12} & 1 - \rho_{11,12}^2 & \rho_{11,21} - \rho_{11,12}\rho_{11,22} \\ \rho_{21,22} - \rho_{21,11}\rho_{11,22} & \rho_{11,21} - \rho_{11,12}\rho_{11,22} & 1 - \rho_{11,22}^2 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{21|11}^2 & \sigma_{21|11}\sigma_{12|11}\rho_{21,12|11} & \sigma_{21|11}\sigma_{22|11}\rho_{21,22|11} \\ \sigma_{21|11}\sigma_{12|11}\rho_{21,12|11} & \sigma_{12|11}^2 & \sigma_{12|11}\sigma_{22|11}\rho_{12,22|11} \\ \sigma_{21|11}\sigma_{22|11}\rho_{21,22|11} & \sigma_{12|11}\sigma_{22|11}\rho_{12,22|11} & \sigma_{22|11}^2 \end{pmatrix}. \end{aligned}$$

Comme l'intérêt est porté sur  $\rho_{21,12|11}$ , le constat est de remarquer que :

$$\sigma_{21|11}\sigma_{12|11}\rho_{21,12|11} = \rho_{21,12} - \rho_{21,11}\rho_{11,12}$$

où

$$\begin{aligned} \sigma_{21|11}^2 &= 1 - \rho_{21,11}^2 \\ \sigma_{12|11}^2 &= 1 - \rho_{11,12}^2 \end{aligned}$$

et donc,

$$\rho_{21,12|11} = \frac{\rho_{21,12} - \rho_{21,11}\rho_{11,12}}{\sqrt{(1 - \rho_{21,11}^2)(1 - \rho_{11,12}^2)}}.$$

Une autre façon d'aborder la situation est de voir que :

$$\psi^{1,2} = \text{diag}(\Psi^{1,2})^{-1/2} \Psi^{1,2} \text{diag}(\Psi^{1,2})^{-1/2}$$

où

$$\text{diag}(\Psi^{1,2})^{-1/2} = \begin{pmatrix} \frac{1}{\sqrt{1-\rho_{21,11}^2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{1-\rho_{12,11}^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-\rho_{22,11}^2}} \end{pmatrix}$$

et d'obtenir  $\rho_{21,12|11} = \psi_{1,2}^{1,2}$ .

Une fois  $\rho_{21,12|11}$  trouvée et  $\psi^{1,2}$  calculée, il est possible d'utiliser  $\psi^{1,2}$  pour calculer  $\psi^{2,2}$  d'où on obtient  $\psi_{1,2}^{2,2} = \rho_{21,22|11,12}$  comme mentionné dans le calcul récursif explicité plus haut.

### 3.3 Recomposition d'une copule gaussienne multivariée à partir de sa décomposition M-Vine

Dans la section précédente, on a montré que le modèle CGM peut être décomposé avec le M-V. Dans cette section, on présente le chemin inverse, c'est-à-dire comment obtenir une CGM à partir de sa décomposition M-Vine, en ayant comme but ultime de construire un modèle VAR(1) à partir d'une décomposition M-V. Pour ce faire, on considère un M-V représenté avec seulement des distributions marginales et des copules bivariées gaussiennes parce que si seulement une copule ou distribution marginale de ce M-V n'est pas gaussienne, il sera impossible de reconstruire une CGM et le modèle sous-jacent pour la série chronologique multivariée ne sera pas un VAR(1).

Considérons un M-V pour un vecteur de séries chronologiques markovien du premier ordre  $k$ -variée tel que toutes ses distributions marginales et ses copules bivariées de la décomposition associée sont gaussiennes, et assumons que les paramètres sont connus. Le modèle M-V est caractérisé par  $d = k(2k - 1)$  copules gaussiennes bivariées. Les paramètres de ces copules gaussiennes représentent les corrélations entre deux processus chronologiques :  $2k - 1$  de ces paramètres correspondent à des corrélations directes provenant du premier arbre de la décomposition, et le reste des paramètres correspondent à des corrélations conditionnelles.

Pour un modèle CGM, les paramètres sont composés de  $d$  corrélations directes entre les séries  $X_{i,t+h-2}$  et  $X_{j,t+l-2}$  avec  $ih \neq jl$ ,  $i, j \in \{1, \dots, k\}$  et  $h, l \in \{1, 2\}$ . Alors, l'objectif est de calculer les  $(k - 1) \times (2k - 1)$  corrélations directes inconnues via les paramètres du M-V. Considérons  $\Psi$ , la matrice contenant tous les paramètres d'une CGM ; voir (3.4) dans la Section 3.2. Les seuls éléments connus de  $\Psi$  sont les  $(2k - 1)$  corrélations directes données par le M-V et positionnées juste au-dessus la diagonale. À partir d'ici, on appelle 'deuxième diagonale' l'ensemble de ces éléments, c'est-à-dire les corrélations directes. L'ensemble des éléments juste au-dessus de la deuxième diagonale s'appelle alors la 'troisième diagonale', et ainsi de suite. Par exemple, la  $(2k)^e$  diagonale de  $\Psi$  est donnée par l'ensemble contenant le seul élément  $\rho_{k1,k2}$  ; la 'première diagonale' (la diagonale classique) est composée de 1.

Pour calculer  $\rho_{ih,jl}$  (la corrélation entre  $X_{i,t+h-2}$  et  $X_{j,t+l-2}$ ) sur la  $q^e$  diagonale de  $\Psi$  pour  $q \geq 3$  (sans perte de généralité  $i > j$ ), considérons la sous-matrice carrée symétrique de  $\Psi$  :

$$\Psi^{(ih,jl)} = \left[ \begin{array}{c|ccc|c} 1 & \rho_{ih,i-1h} & \cdots & \rho_{ih,j+1l} & \rho_{ih,jl} \\ \hline \rho_{ih,i-1h} & 1 & \cdots & \rho_{i-1h,j+1l} & \rho_{i-1h,jl} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline \rho_{ih,j+1l} & \rho_{i-1h,j+1l} & \cdots & 1 & \rho_{j+1l,jl} \\ \hline \rho_{ih,jl} & \rho_{i-1h,jl} & \cdots & \rho_{j+1l,jl} & 1 \end{array} \right] = \left[ \begin{array}{c|c|c} 1 & \Psi_{21}^{(1)} & \rho_{ih,jl} \\ \hline \Psi_{12}^{(1)} & \Psi_{22} & \Psi_{12}^{(2)} \\ \hline \rho_{ih,jl} & \Psi_{21}^{(2)} & 1 \end{array} \right]$$

On définit :

$$\begin{aligned} \Psi_1^{(ih,jl)} &= \begin{pmatrix} 1 & \rho_{ih,jl} \\ \rho_{ih,jl} & 1 \end{pmatrix}, & \Psi_2^{(ih,jl)} &= \Psi_{22}, \\ \Psi_3^{(ih,jl)} &= \begin{pmatrix} \Psi_{21}^{(1)} \\ \Psi_{21}^{(2)} \end{pmatrix}, & \Psi_4^{(ih,jl)} &= \begin{pmatrix} \Psi_{12}^{(1)} & \Psi_{12}^{(2)} \end{pmatrix}. \end{aligned}$$

Prendre note que le partitionnement de  $\Psi^{(ih,jl)}$  ci-haut est celui pour le cas où  $h, l = 1$ . Pour  $h, l = 2$ , la sous-matrice à considérer est donnée par  $(\Psi^{(ih,jl)})^{T'}$ . Pour  $h = 1$  et  $l = 2$  la sous-matrice  $\Psi^{(ih,jl)}$  de  $\Psi$  est plutôt donnée comme suit :

$$\left( \begin{array}{c|cccccc|c} 1 & \rho_{ih,i-1h} & \cdots & \rho_{ih,1h} & \rho_{ih,1l} & \cdots & \rho_{ih,j-1l} & \rho_{ih,jl} \\ \hline \rho_{ih,i-1h} & 1 & \cdots & \rho_{i-1h,1h} & \rho_{i-1h,1l} & \cdots & \rho_{i-1h,j-1l} & \rho_{i-1h,jl} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \hline \rho_{ih,1h} & \rho_{i-1h,1h} & \cdots & 1 & \rho_{1h,1l} & \cdots & \rho_{1h,j-1l} & \rho_{1h,jl} \\ \rho_{ih,1l} & \rho_{i-1h,1l} & \cdots & \rho_{1h,1l} & 1 & \cdots & \rho_{1l,j-1l} & \rho_{1l,jl} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline \rho_{ih,j-1l} & \rho_{i-1h,j-1l} & \cdots & \rho_{1h,j-1l} & \rho_{1l,j-1l} & \cdots & 1 & \rho_{j-1l,jl} \\ \hline \rho_{ih,jl} & \rho_{i-1h,jl} & \cdots & \rho_{1h,jl} & \rho_{1l,jl} & \cdots & \rho_{j-1l,jl} & 1 \end{array} \right).$$

Dans le cas où  $h = 2$  et  $l = 1$ , la matrice d'intérêt est la transposée de cette dernière.

Les prochains développements prennent en considération que nous sommes dans le cas où  $i > j$  et que  $h = l = 1$ . Pour les autres cas, on a exactement les mêmes développements avec seulement des différences mineures dans les indices. Continuons.

Comme dans la section 3.2, on suppose que les distributions marginales sont des gaussiennes centrées et réduites pour le calcul des corrélations. Alors,  $\Psi^{(ih,jl)}$  représente la matrice de variance-covariance de  $(X_{i;t+h-2}, X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}, X_{j;t+l-2})^T$ , qui se

distribue selon une loi gaussienne multivariée centrée et réduite. De plus, pour calculer les paramètres sur la  $q^e$  diagonale, considérons que tous les paramètres dans  $\Psi$  jusqu'à la  $(q-1)^e$  diagonale sont connus. De (3.5), il s'ensuit que

$$\begin{aligned} \begin{bmatrix} \sigma_{ih|D}^2 & \rho_{ih,jl|D}\sigma_{ih|D}\sigma_{jl|D} \\ \rho_{ih,jl|D}\sigma_{ih|D}\sigma_{jl|D} & \sigma_{jl|D}^2 \end{bmatrix} &= \Psi_1^{(ih,jl)} - \Psi_3^{(ih,jl)} \left( \Psi_2^{(ih,jl)} \right)^{-1} \Psi_4^{(ih,jl)} \\ &= \begin{bmatrix} 1 & \rho_{ih,jl} \\ \rho_{ih,jl} & 1 \end{bmatrix} - \beta^{(ih,jl)} \\ &= \begin{bmatrix} 1 - \beta_{11}^{(ih,jl)} & \rho_{ih,jl} - \beta_{12}^{(ih,jl)} \\ \rho_{ih,jl} - \beta_{12}^{(ih,jl)} & 1 - \beta_{22}^{(ih,jl)} \end{bmatrix} \end{aligned} \quad (3.7)$$

où  $D = \{(i-1)h, \dots, (j+1)l\}$ ,  $\sigma_{ih|D}^2$  et  $\sigma_{jl|D}^2$  sont les variances conditionnelles des variables  $X_{i;t+h-2}|X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}$  et  $X_{j;t+l-2}|X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}$ , respectivement ;  $\rho_{ih,jl|D}$  est l'un des  $2k - q + 1$  paramètres du  $q^e$  arbre du M-V et  $\beta^{(ih,jl)}$  est une matrice de format  $2 \times 2$  qui peut s'écrire

$$\beta^{(ih,jl)} = \Psi_3^{(ih,jl)} \left( \Psi_2^{(ih,jl)} \right)^{-1} \Psi_4^{(ih,jl)} = \begin{pmatrix} \beta_{11}^{(ih,jl)} & \beta_{12}^{(ih,jl)} \\ \beta_{21}^{(ih,jl)} & \beta_{22}^{(ih,jl)} \end{pmatrix}$$

avec  $\beta_{12}^{(ih,jl)} = \beta_{21}^{(ih,jl)}$ .

L'équation (3.7) mène à un système de trois équations avec trois inconnues :

$$\begin{aligned} \sigma_{ih|D}^2 &= 1 - \beta_{11}^{(ih,jl)} \\ \sigma_{jl|D}^2 &= 1 - \beta_{22}^{(ih,jl)} \\ \rho_{ih,jl|D}\sigma_{ih|D}\sigma_{jl|D} &= \rho_{ih,jl} - \beta_{12}^{(ih,jl)}. \end{aligned} \quad (3.8)$$

Depuis (3.8), on obtient :

$$\rho_{ih,jl} = \rho_{ih,jl|D} \sqrt{\left(1 - \beta_{11}^{(ih,jl)}\right) \left(1 - \beta_{22}^{(ih,jl)}\right)} + \beta_{12}^{(ih,jl)}. \quad (3.9)$$

Ainsi, toutes les corrélations directes de  $\Psi$  peuvent être obtenues à partir des paramètres d'un M-V. La procédure récursive ci-dessous résume les développements qui fournissent



tous les paramètres d'une CGM à partir d'un M-V avec des distributions marginales et copules bivariées gaussiennes :

1. Soit  $\Psi$  la matrice symétrique de dimension  $2k \times 2k$  contenant les paramètres de la CGM partitionnée comme en (3.4) et considérons  $\Omega'$  l'ensemble des paramètres du M-V.
2. Les éléments sur la deuxième diagonale de  $\Psi$  sont obtenus depuis les paramètres du premier arbre du M-V dans  $\Omega'$ .
3. Pour  $q$  allant de 3 jusqu'à  $2k$ ,
  - (a) Considérer la  $q^e$  diagonale de  $\Psi$  composée de  $2k - q + 1$  paramètres.
  - (b) Pour  $u$  allant de 1 jusqu'à  $2k - q + 1$ ,
    - i. Considérer le  $u^e$  élément sur la  $q^e$  diagonale de  $\Psi$  qui correspond à un paramètre inconnu  $\rho_{ih,jl}$  avec :
 
$$\begin{cases} i = k - u + 1 & \text{et } h = 1 & \text{si } u \leq k \\ i = u - k & \text{et } h = 2 & \text{sinon} \end{cases}$$
 et
 
$$\begin{cases} j = k - u - q + 2 & \text{et } l = 1 & \text{si } q + u \leq k + 1 \\ j = q + u - k - 1 & \text{et } l = 2 & \text{sinon.} \end{cases}$$
    - ii. Calculer  $\beta^{(ih,jl)}$  comme décrit plus haut.
    - iii. En utilisant (3.9) et  $\rho_{ih,jl|i-1h,\dots,j+1l} \in \Omega'$ , calculer  $\rho_{ih,jl}$ .
    - iv. Entrer la valeur trouvée pour  $\rho_{ih,jl}$  dans  $\Psi$ .
4. Toutes les entrées de la matrice  $\Psi$  sont calculées.

Comme pour la section précédente sur le passage d'une CGM à un M-V, illustrons ces derniers développements à l'aide d'une série chronologique bivariée  $\mathbf{X}_t = (X_{1;t}, X_{2;t})$  markovienne d'ordre 1 dont la dépendance est représentée à l'aide d'une décomposition M-Vine gaussienne avec paramètres donnés. Ainsi, on connaît  $\rho_{21,11}$ ,  $\rho_{11,12}$ ,  $\rho_{21,12|11}$ ,

$\rho_{11,22|12}$  et  $\rho_{21,22|11,12}$ . On considère la matrice qui regroupe toutes les corrélations directes nécessaires à une représentation CGM suivante :

$$\Psi = \begin{pmatrix} 1 & \rho_{21,11} & \rho_{21,12} & \rho_{21,22} \\ \rho_{21,11} & 1 & \rho_{11,12} & \rho_{11,22} \\ \rho_{21,12} & \rho_{11,12} & 1 & \rho_{11,21} \\ \rho_{21,22} & \rho_{11,22} & \rho_{11,21} & 1 \end{pmatrix}$$

dont on connaît seulement les termes au-dessus et en-dessous de la diagonale.

Ainsi, il faut trouver les corrélations  $\rho_{21,12}$ ,  $\rho_{11,22}$  et  $\rho_{21,22}$ . Comme expliqué plus haut, il faut commencer par le premier terme de la troisième diagonale, c'est-à-dire  $\rho_{21,12}$ . On considère donc la sous-matrice

$$\Psi^{(21,12)} = \left( \begin{array}{c|c|c} 1 & \rho_{21,11} & \rho_{21,12} \\ \hline \rho_{21,11} & 1 & \rho_{11,12} \\ \hline \rho_{21,12} & \rho_{11,12} & 1 \end{array} \right).$$

On pose alors :

$$\begin{aligned} \Psi_1^{(21,12)} &= \begin{pmatrix} 1 & \rho_{21,12} \\ \rho_{21,12} & 1 \end{pmatrix}, & \Psi_2^{(21,12)} &= 1, \\ \Psi_3^{(21,12)} &= \begin{pmatrix} \rho_{21,11} \\ \rho_{11,12} \end{pmatrix}, & \Psi_4^{(21,12)} &= \begin{pmatrix} \rho_{21,11} & \rho_{11,12} \end{pmatrix}. \end{aligned}$$

et on obtient :

$$\beta^{(21,12)} = \Psi_3^{(21,12)} \left( \Psi_2^{(21,12)} \right)^{-1} \Psi_4^{(21,12)} = \begin{pmatrix} \rho_{21,11}^2 & \rho_{21,11}\rho_{11,12} \\ \rho_{21,11}\rho_{11,12} & \rho_{11,12}^2 \end{pmatrix}.$$

Finalement,

$$\rho_{21,12} = \rho_{21,12|11} \sqrt{(1 - \rho_{21,11}^2)(1 - \rho_{11,12}^2)} + \rho_{21,11}\rho_{11,12}.$$

Ensuite, de la même façon, il est possible d'obtenir que

$$\rho_{11,22} = \rho_{11,22|12} \sqrt{(1 - \rho_{21,11}^2)(1 - \rho_{11,12}^2)} + \rho_{21,11}\rho_{11,12}.$$

En connaissant toutes les valeurs de la troisième diagonale de  $\Psi$ , il est possible de faire le même travail pour la quatrième diagonale. Cependant, cette fois-ci la séparation de  $\Psi$  est donnée par :

$$\Psi^{(21,22)} = \left( \begin{array}{c|cc|c} 1 & \rho_{21,11} & \rho_{21,12} & \rho_{21,22} \\ \hline \rho_{21,11} & 1 & \rho_{11,12} & \rho_{11,22} \\ \rho_{21,12} & \rho_{11,12} & 1 & \rho_{11,21} \\ \hline \rho_{21,22} & \rho_{11,22} & \rho_{11,21} & 1 \end{array} \right).$$

### 3.4 Retour au modèle classique VAR(1) à partir d'une copule gaussienne multivariée

Dans les sections précédentes, on a étudié la représentation CGM d'un VAR(1), la décomposition M-V d'un modèle CGM et on a reconstruit une CGM à partir de sa décomposition M-V. Dans cette section, l'objectif est de trouver le processus VAR(1) sous-jacent à une CGM donnée. Soit  $\mathbf{X}_t = (X_{1;t}, X_{2;t}, \dots, X_{k;t})^T$  un processus markovien de premier ordre  $k$ -varié représenté par une CGM, c'est-à-dire que les  $k$  distributions marginales gaussiennes et tous les paramètres, qui sont les corrélations directes au même temps ( $t$ ) et aux temps croisés ( $t$  et  $t-1$ ), sont identifiés. La moyenne et la variance de  $X_{i;t+h-2}$  sont respectivement notées  $\mu_i$  et  $\sigma_i^2$ , avec  $i \in \{1, \dots, k\}$  et  $h \in \{1, 2\}$ . Alors, la distribution conjointe de  $(\mathbf{X}_{t-1}^T, \mathbf{X}_t^T)^T$  est donnée par :

$$\begin{pmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_t \end{pmatrix} \sim N \left( \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2^T & \boldsymbol{\Sigma}_1 \end{pmatrix} \right) \quad (3.10)$$

où

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}, \boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_1^2 & \rho_{11,21}\sigma_1\sigma_2 & \cdots & \rho_{11,k1}\sigma_1\sigma_k \\ \rho_{11,21}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{21,k1}\sigma_2\sigma_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{11,k1}\sigma_1\sigma_k & \rho_{21,k1}\sigma_2\sigma_k & \cdots & \sigma_k^2 \end{pmatrix},$$

et

$$\Sigma_2 = \begin{pmatrix} \rho_{11,12}\sigma_1^2 & \rho_{11,22}\sigma_1\sigma_2 & \cdots & \rho_{11,k2}\sigma_1\sigma_k \\ \rho_{21,12}\sigma_1\sigma_2 & \rho_{21,22}\sigma_2^2 & \cdots & \rho_{21,k2}\sigma_2\sigma_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1,12}\sigma_1\sigma_k & \rho_{k1,22}\sigma_2\sigma_k & \cdots & \rho_{k1,k2}\sigma_k^2 \end{pmatrix}.$$

Pour obtenir le modèle VAR(1) sous-jacent, on a besoin de la distribution de  $\mathbf{X}_t|\mathbf{X}_{t-1}$ .

À partir de (3.10) et de la théorie conditionnelle pour les distributions gaussiennes :

$$\begin{aligned} \boldsymbol{\mu}_{\mathbf{X}_t|\mathbf{X}_{t-1}} &= \boldsymbol{\mu} + \Sigma_2^T \Sigma_1^{-1} (\mathbf{X}_{t-1} - \boldsymbol{\mu}) \\ &= (\mathbb{I} - \Sigma_2^T \Sigma_1^{-1}) \boldsymbol{\mu} + \Sigma_2^T \Sigma_1^{-1} \mathbf{X}_{t-1} \\ &= \mathbf{v} + \Phi \mathbf{X}_{t-1} \end{aligned} \tag{3.11}$$

et

$$\Sigma_{\mathbf{X}_t|\mathbf{X}_{t-1}} = \Sigma_1 - \Sigma_2^T \Sigma_1^{-1} \Sigma_2 = \bar{\Sigma}. \tag{3.12}$$

Alors, avec (3.11), (3.12),  $\mathbf{v} = (\mathbb{I} - \Sigma_2^T \Sigma_1^{-1}) \boldsymbol{\mu}$ ,  $\Phi = \Sigma_2^T \Sigma_1^{-1}$  et  $\bar{\Sigma} = \Sigma_1 - \Sigma_2^T \Sigma_1^{-1} \Sigma_2$ , on obtient que  $\mathbf{X}_t|\mathbf{X}_{t-1} \sim N(\mathbf{v} + \Phi \mathbf{X}_{t-1}, \bar{\Sigma})$ , et on trouve l'expression VAR(1) pour  $\mathbf{X}_t$  :

$$\mathbf{X}_t = \mathbf{v} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$$

où  $\boldsymbol{\epsilon}_t \sim N(0, \bar{\Sigma})$ .

On en conclut qu'un processus chronologique représenté avec une CGM peut être réécrit comme un VAR(1). Aussi, on note qu'il n'y a aucune condition requise sur la CGM pour faire le changement de représentation. Cela implique que toutes les CGM ont un VAR(1) équivalent. Ceci ferme la boucle (Figure 3.1) initiée à la section 3.1.

# CHAPITRE 4

## Applications

Le lien et les passages entre un processus VAR(1) et une décomposition M-V ayant été soulignés dans le précédent chapitre, on s'intéresse ici à une application de cette méthodologie sur un exemple purement théorique tirée de l'article de Beare et Seo (2015) [BS15] (Section 4.1) et à un autre empirique portant sur la fluctuation du prix du blé dans trois régions du monde (Section 4.2).

La première application est théorique et fait intervenir directement les calculs récurrents détaillés précédemment pour obtenir une décomposition M-V à partir d'un processus VAR(1). L'exemple est tiré de l'article de Beare et Seo (2015) [BS15] et a été utilisé dans cet article pour illustrer que le M-V est une généralisation du VAR(1). Nous vérifions l'exactitude de la décomposition M-V proposée par Beare et Seo (2015) [BS15]. Le passage de la décomposition M-V purement gaussienne vers un processus VAR(1) sera aussi présenté à l'aide d'un exemple donné.

La deuxième partie de ce chapitre porte sur un jeu de données sur le prix du blé observé mensuellement entre Novembre 2010 et Avril 2014 dans trois pays : Australie (Aus), Kazakhstan (Kaz) et Royaume-Uni (RU). Ce jeu de données provient du site officiel de "Food

and Agriculture Organization" (FAO). On s'intéressera à l'estimation des paramètres d'une décomposition M-V gaussienne de cette série chronologique tridimensionnelle dans l'optique de faire une transition vers un processus VAR(1). Une décomposition M-V avec plusieurs choix de copules sera aussi considérée.

## 4.1 Application théorique

Beare et Seo (2015) [BS15] discute d'une relation entre VAR(1) et M-V. Sans explications, ils suggèrent qu'une série chronologique bivariée markovienne du premier ordre représentée par le modèle VAR(1)  $\mathbf{X}_t = \mathbf{v} + \phi\mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$  où  $\boldsymbol{\epsilon}_t \sim i.i.d.N(0, \boldsymbol{\Sigma})$  et les paramètres sont :

$$\mathbf{v} = \begin{bmatrix} 3.5 \\ -0.5 \end{bmatrix}, \quad \phi = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \quad (4.1)$$

est équivalent à une décomposition M-V où les distribution marginales sont gaussiennes données par  $N(5, 1.6638^2)$  pour  $X_{1t}$  et par  $N(0, 1.2377^2)$  pour  $X_{2t}$  et les paramètres étant :

$$\begin{aligned} \rho_{11,21} &= 0.7243, & \rho_{11,12} &= 0.4078, & \rho_{12,22} &= 0.7243, \\ \rho_{21,12|11} &= 0.1123, & \rho_{11,22|12} &= -0.1406, & \rho_{21,22|11,12} &= -0.0144. \end{aligned} \quad (4.2)$$

On veut revoir cet exemple et vérifier si l'équivalence suggérée est valide. Cette vérification est faite en considérant le modèle VAR(1) (4.1) et en changeant sa représentation de VAR(1) vers CGM (voir Section 3.1) et ensuite de CGM vers M-V (voir Section 3.2). Inversement, en débutant avec la décomposition M-V (4.2) de Beare et Seo (2015) [BS15], on obtiendra le modèle VAR(1) équivalent via une CGM (voir Section 3.3 et Section 3.4)

Commençons par considérer le VAR(1) donné par (4.1). Après calcul, un modèle CGM est obtenu avec les marginales  $X_{1t} \sim N(5, 1.1053^2)$  et  $X_{2t} \sim N(0, 1.017^2)$  et les paramètres

suivants :

$$\begin{aligned}\rho_{11,21} &= 0.5219, & \rho_{11,12} &= 0.3960, & \rho_{12,22} &= 0.5219, \\ \rho_{21,12} &= 0.3406, & \rho_{11,22} &= 0.1609, & \rho_{21,22} &= 0.1567.\end{aligned}\tag{4.3}$$

À partir de ce modèle CGM (4.3), on trouve la décomposition M-V avec les mêmes marginales  $X_{1t} \sim N(5, 1.1053^2)$  et  $X_{2t} \sim N(0, 1.017^2)$  et les paramètres

$$\begin{aligned}\rho_{11,21} &= 0.5219, & \rho_{11,12} &= 0.3960, & \rho_{12,22} &= 0.5219, \\ \rho_{21,12|11} &= 0.1709, & \rho_{11,22|12} &= -0.0585, & \rho_{21,22|11,12} &\approx 0\end{aligned}\tag{4.4}$$

qui n'est pas la décomposition M-V (4.2) proposée par Beare et Seo (2015) [BS15].

Pour faire une vérification consciencieuse, considérons la décomposition M-V (4.2) proposée par Beare et Seo (2015) [BS15]. À partir de cette décomposition, on obtient une CGM avec les marginales  $N(5, 1.6638^2)$  pour  $X_{1t}$  et  $N(0, 1.2377^2)$  pour  $X_{2t}$  et les paramètres suivants :

$$\begin{aligned}\rho_{11,21} &= 0.7243, & \rho_{11,12} &= 0.4078, & \rho_{12,22} &= 0.7243, \\ \rho_{21,12} &= 0.3661, & \rho_{11,22} &= 0.2068, & \rho_{21,22} &= 0.1974.\end{aligned}\tag{4.5}$$

De ce modèle CGM (4.5), on obtient un modèle VAR(1) avec paramètres :

$$\mu = \begin{bmatrix} 3.5 \\ -0.5 \end{bmatrix}, \quad \phi = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2.2787 & 1.3033 \\ 1.3033 & 1.4591 \end{bmatrix},\tag{4.6}$$

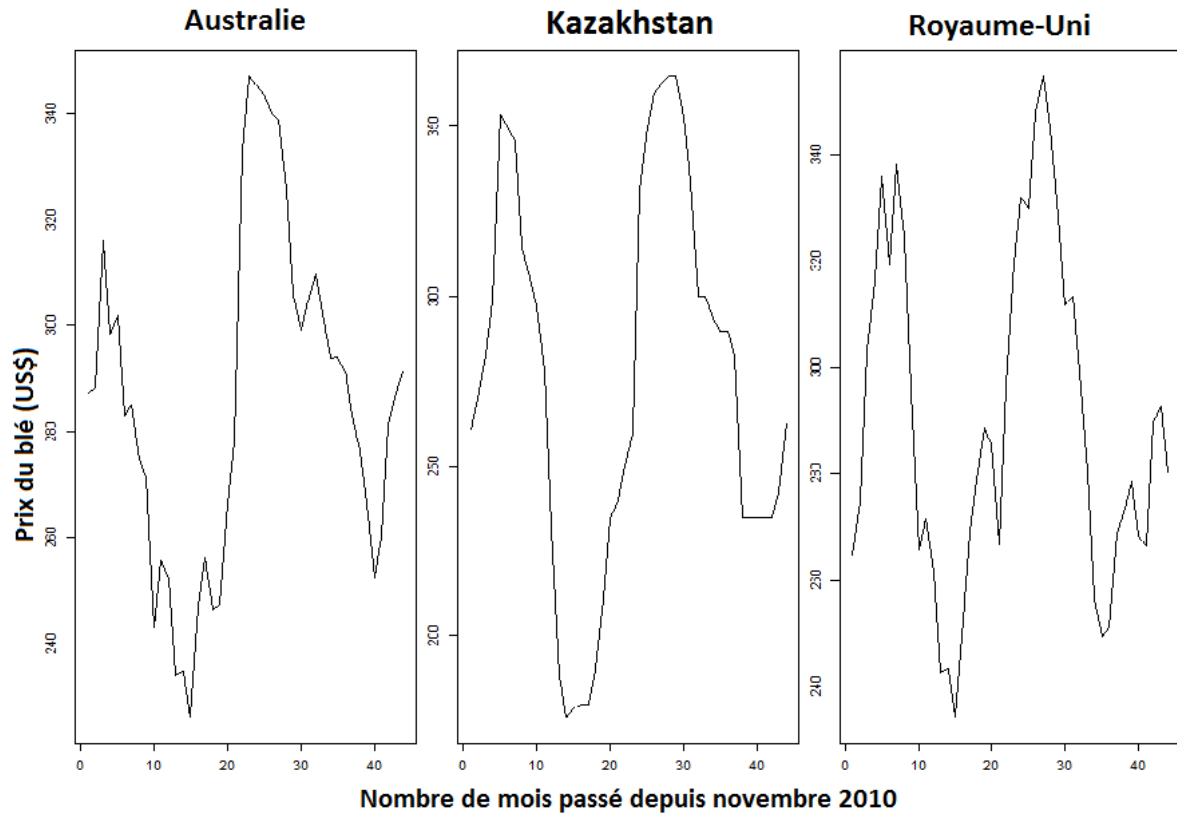
qui n'est pas le modèle VAR(1) (4.1) introduit par Beare et Seo (2015) [BS15].

En conclusion, les calculs derrière le changement de représentation de VAR(1) vers M-V fait par Beare et Seo (2015) [BS15] ne sont pas corrects. En fait, les auteurs ont obtenu (4.2) alors que le vrai modèle M-V équivalent est (4.4) donné plus haut. On observe aussi que les paramètres du modèle VAR(1) équivalent à la décomposition M-V proposé par Beare et Seo (2015) [BS15] [voir (4.6) ci-dessus] sont différents de ceux du modèle VAR(1) décrit dans le même article [(4.1) ci-dessus].

## 4.2 Application empirique : prix du blé

Le jeu de données utilisé ici pour cette application empirique provient du site officiel de *Food and Agriculture Organization* (FAO). Il est composé des prix du blé observés mensuellement de novembre 2010 à avril 2014 (42 observations) dans trois pays : L'Australie (Aus), le Kazakhstan (Kaz) et le Royaume-Uni (UK).

Figure 4.1 – Séries chronologiques des prix du blé entre novembre 2010 et avril 2014 en Australie, au Kazakhstan et au Royaume-Uni



Pour savoir la meilleure décomposition M-Vine pour une série chronologique donnée, on doit d'abord considérer tous les arrangements possibles de copules bivariées en prenant chaque copule bivariée à partir d'une sélection de famille prédéfinie, ensuite maximiser la vraisemblance en variant chacun des paramètres de chacune des copules pour chacun



des arrangements et finalement retenir l'arrangement qui fournit la plus grande valeur de vraisemblance totale. En considérant  $p$  familles de copules pour une série chronologique  $k$ -variée, il y a  $(k/2(3k-1))^p$  arrangements possibles. Alors, avec des valeurs de  $p$  et de  $k$  modérément élevées le nombre d'arrangements possibles monte très rapidement. En effet, en prenant une série chronologique tri-variée et en choisissant les copules à partir d'une petite sélection, disons 6 familles, on obtient près de 3 millions d'arrangements possibles (2 985 984 arrangements pour être exact). C'est la raison pour laquelle l'utilisation d'une autre méthode d'estimation des paramètres d'une décomposition M-V est nécessaire.

En vue de palier à ce problème, l'estimation semi-paramétrique de modèle à base de copules dans le cadre univarié a déjà été introduite par Chen et Fan (2006) [CF06] et étendue, pour le cas multivarié, par Rémillard (2012) [RPS12]. En bref, dans un cas  $k$ -varié, leur approche consiste à estimer les  $k$  marginales du modèle de manière non paramétrique pour ensuite estimer les copules et leur paramètre de manière paramétrique à l'aide des marginales obtenues.

Beare et Seo (2015) [BS15] s'inspire de cette approche pour utiliser une méthode d'estimation semi-paramétrique. Cette procédure, qui s'appelle "Stepwise Semiparametric" (SSP), a été tout d'abord proposée par Aas et collaborateurs (2009) [ACFB09] et étudiée en détail par Hobaek et Haff en 2012 [HH12] et en 2013 [HH13]. Elle consiste à estimer chaque arbre les uns après les autres par la méthode du maximum de vraisemblance et à considérer que la meilleure estimation pour un arbre conduira à une décomposition finale de M-Vine proche de la meilleure estimation globale qui serait obtenue en utilisant tous les arrangements possibles. Pour de plus amples détails sur cette méthode d'estimation semi-paramétrique basée sur le maximum de vraisemblance, se référer à Beare et Seo (2015) [BS15].

Il est important de se rappeler qu'un VAR(1) peut être représenté comme un modèle M-V si et seulement si toutes les distributions marginales et toutes les copules bivariées

de la décomposition M-V sont gaussiennes. Dans un exemple empirique comme celui-ci sur les prix du blé, un mélange de différentes copules peut être attendu. Néanmoins, la méthode SSP est utilisée tout comme dans l'article de Beare et Seo (2015) [BS15], sauf qu'on considère d'abord seulement la famille des copules gaussiennes. Ainsi, on obtient le mélange de copules gaussiennes bivariées qui maximise la vraisemblance. Ensuite, la représentation du modèle M-V gaussien peut être changée pour un VAR(1) si les distributions marginales sont, ou peuvent raisonnablement être considérées, gaussiennes.

Dans notre application empirique, l'ordre des trois séries chronologiques des prix du blé influence le maximum de vraisemblance obtenu, puisque les copules qui sont estimées varient avec l'ordre dans lequel on considère les séries. Alors, il est nécessaire d'effectuer l'estimation pour la méthode SSP pour chaque arrangement possible des trois séries (6 arrangements possibles) et de retenir l'arrangement pour lequel le maximum de vraisemblance obtenu est le plus élevé. Ainsi, en considérant la méthode d'estimation SSP et seulement la famille de copules gaussiennes on obtient que l'ordre optimal est Aus-Kaz-UK. En conséquence, 11 et 12 feront référence aux prix du blé en Australie aux temps  $(t - 1)$  et  $t$  respectivement, pareillement pour 21 et 22 pour le Kazakhstan et 31 et 32 pour le Royaume-Uni. Dans le tableau 4.1, les résultats de l'estimation pour le meilleur mélange de copules gaussiennes sont détaillés.

D'après le tableau 4.1, on note tout d'abord une association positive entre les valeurs de maximum de vraisemblance et la force de la structure de dépendance dans le temps. En fait, les plus grandes valeurs de maximum de vraisemblance ( $>10$ ) proviennent du premier arbre (31,21 (32,22), 21,11 (22,12) et 11,12) et des copules correspondant à la dépendance conditionnelle dans une même série (21,22|11,12 et 31,32|21,11,12,22) pour lesquelles les valeurs des paramètres estimées sont plus élevées que 0,65. Comme une copule bivariée indépendante a une contribution théorique nulle à la vraisemblance, il semble que les copules bivariées 21,12|11, 31,12|21,11 et 11,32|12,22 peuvent être considérées presque

Tableau 4.1 – Paramètres estimés pour une représentation M-V des prix du blé utilisant des copules gaussiennes

Arbre	Copule	Paramètre estimé	Maximum de vraisemblance
1	31,21 (32,22)	0.797	38.557
	21,11 (22,12)	0.748	30.814
	11,12	0.896	32.106
2	31,11 21 (32,12 22)	0.408	7.375
	21,12 11	-0.144	0.428
	11,22 12	0.594	9.037
3	31,12 21,11	-0.089	0.169
	21,22 11,12	0.849	26.900
	11,32 12,22	-0.174	0.842
4	31,22 21,11,12	0.426	4.291
	21,32 11,12,22	-0.197	1.057
5	31,32 21,11,12,22	0.659	13.930
Maximum de vraisemblance total			165.50

indépendantes.

Deuxièmement, de grandes valeurs de maximum de vraisemblance ont été obtenues pour 21,22|11,12 et 31,32|21,11,12,22. Pour 21,22|11,12, on rappelle que la deuxième série des prix du blé est celle du Kazakhstan. Concrètement, le prix du blé au Kazakhstan d'un mois passé affecte seulement le prix du blé au Kazakhstan dans le mois en cours ; les prix du blé pour l'Australie et le Royaume-Uni dans le mois en cours ne sont pas influencés par le prix du blé du Kazakhstan dans le mois passé, donc pas d'effet de causalité.

Finalement, comme le travail est effectué sur le prix du blé d'une même récolte dans trois parties du monde, on peut s'attendre que toutes les structures de dépendence entre les prix du blé à ces endroits soient positives. Toutefois, on note quatre corrélations négatives dans le tableau 4.1 desquels 21,32|11,12,22 possède la plus grande valeur de maximum de vraisemblance et de paramètre estimé en valeur absolue. Ceci est lié aux copules

conditionnelles, c'est-à-dire aux corrélations conditionnelles, et leur valeur négative est le résultat du conditionnement. En fait, la corrélation directe entre le Kazakhstan au temps  $(t - 1)$  (21) et le Royaume-Uni au temps  $t$  (32), qui n'est pas rapportée dans le tableau 4.1, est plus grande que 0,6 alors que sa valeur conditionnelle dans le tableau 4.1 est -0,197. Des observations similaires peuvent être effectuées pour les trois autres copules gaussiennes conditionnelles avec paramètre estimé négatif.

D'après le test de normalité de Kolmogorov-Smirnov, qui est plus robuste que le test de normalité de Shapiro-Wilk pour des autocorrélations temporelles selon Dutilleul et Legendre (1992) [DL92], il n'est pas déraisonnable d'assumer que les trois séries chronologiques des prix du blé suivent des distributions normales. Donc, il est justifié d'utiliser les notions abordées dans le chapitre 3 puisqu'on a une décomposition M-V purement gaussienne avec des distributions marginales gaussiennes. Ainsi, le passage d'une représentation M-V à VAR(1) via une CGM est possible en effectuant successivement le calcul récursif de la section 3.3 et les opérations de la section 3.4. Ceci pourrait être un nouveau moyen d'estimer les paramètres d'un VAR(1) pour des données sur des séries chronologiques.

La moyenne et la variance de chaque série sont estimées par :  $\hat{\mu}_1 = 285.55$ ,  $\hat{\sigma}_1^2 = 1036.8$ ,  $\hat{\mu}_2 = 275.8$ ,  $\hat{\sigma}_2^2 = 3326.7$ ,  $\hat{\mu}_3 = 290.05$  et  $\hat{\sigma}_3^2 = 1021.85$ . Avec ces moyennes et variances estimées (pour les distributions marginales gaussiennes) et les paramètres estimés pour les copules bivariées gaussiennes dans le tableau 4.1, la représentation M-V peut être réécrite comme une représentation CGM avec une matrice

$$\Psi = \begin{pmatrix} 1 & 0,7482 & 0,7601 & 0,8960 & 0,8735 & 0,7055 \\ 0,7482 & 1 & 0,7972 & 0,6280 & 0,9212 & 0,6003 \\ 0,7601 & 0,7972 & 1 & 0,6374 & 0,8583 & 0,8044 \\ 0,8960 & 0,6280 & 0,6374 & 1 & 0,7972 & 0,7601 \\ 0,8735 & 0,9212 & 0,8583 & 0,7972 & 1 & 0,7482 \\ 0,7055 & 0,6003 & 0,8044 & 0,7601 & 0,7482 & 1 \end{pmatrix}$$

(cette matrice est construite en considérant l'ordre 31,21,11,12,22,32), d'où on tire une

représentation VAR(1)  $\mathbf{X}_t = \mathbf{v} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$ , où  $\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma})$  et les paramètres sont :

$$\mathbf{v} = \begin{pmatrix} 31.471 \\ -140.645 \\ 19.826 \end{pmatrix}, \quad \Phi = \begin{pmatrix} 0.9948 & -0.0333 & -0.0717 \\ 0.6282 & 0.5134 & 0.3291 \\ 0.2979 & -0.1272 & 0.7593 \end{pmatrix},$$

$$\text{et } \boldsymbol{\Sigma} = \begin{pmatrix} 198.50 & -8.70 & 155.80 \\ -8.70 & 213.26 & 174.92 \\ 155.80 & 174.92 & 322.19 \end{pmatrix}.$$

Une question fondamentale est alors soulevée : est-ce qu'une décomposition M-Vine vaut le travail supplémentaire nécessaire ou un modèle VAR(1) est suffisant pour ce jeu de données ? Pour répondre à cette question, considérons encore la méthode d'estimation SSP pour les paramètres du M-V, mais appliquée en utilisant des copules provenant de plusieurs familles préselectionnées : indépendante, gaussienne, Student, Clayton, Frank, Gumbel et Ali-Mikhail-Haq. Il est très possible que le modèle M-V estimé soit composé de copules de plusieurs familles différentes cette fois-ci.

En considérant les sept familles de copules mentionnées, l'ordre optimal des séries est complètement renversé comparé à ce que l'on avait obtenu en étudiant le cas avec seulement la famille des copules gaussiennes (Tableau 4.1) : on a désormais l'ordre optimal UK-Kaz-Aus. La notation utilisée plus tôt ne change pas et la notation est : 31 et 32 pour le Royaume-Uni, 21 et 22 pour le Kazakhstan et 11 et 12 pour l'Australie. Les résultats obtenus pour la représentation M-V sont donnés dans le tableau 4.2.

Plus le maximum de vraisemblance est élevé, meilleur est le modèle M-V estimé. On observe une augmentation de 1.05 du maximum de vraisemblance total entre le mélange de copules gaussiennes et le mélange de copules de plusieurs familles différentes. Une telle augmentation (de <1% relativement parlant) suggère que le M-V estimé à partir d'un mixe de différentes copules est légèrement meilleur que le M-V estimé à partir de seulement des copules gaussiennes dans ce cas-ci. Comme le M-V estimé depuis des copules gaussiennes bivariées est équivalent à un modèle VAR(1) (sous la condition que

Tableau 4.2 – Résultats pour une représentation M-V des prix du blé en utilisant des copules à partir d'un ensemble de familles préselectionnées

Arbre	Copule	Famille de copule estimé	Paramètre estimé	Maximum de vraisemblance
1	11,21 (12,22)	Clayton	2.351	38.637
	21,31 (22,32)	Gaussienne	0.797	38.557
	31,32	Gaussienne	0.862	26.555
2	11,31 21 (12,32 22)	Gumbel	1.489	11.202
	21,32 31	Clayton	-0.061	0.053
	31,22 32	Gumbel	1.590	8.413
3	11,32 21,31	Frank	1.265	1.049
	21,22 31,32	Gaussienne	0.821	23.816
	31,12 32,22	Gaussienne	-0.397	3.927
4	11,22 21,31,32	Gumbel	1.465	6.069
	21,12 31,32,22	Clayton	-0.132	0.524
5	11,12 21,31,32,22	Student	0.464	7.746
Maximum de vraisemblance total				166.55

les distributions marginales peuvent raisonnablement être considérées gaussiennes), un M-V avec des copules bivariées différentes estimé avec la méthode SSP ne représente pas une grande amélioration par rapport au modèle VAR(1) dans cet exemple empirique.

# CONCLUSION

Beare et Seo (2015) [BS15] ont proposé la décomposition M-Vine afin de représenter des processus pour séries chronologiques multivariées. Dans ce mémoire, une clarification de certains de leurs développements concernant la relation entre un modèle VAR (1) et sa décomposition M-Vine, via une copule gaussienne multivariée et inversement, à partir de VAR(1) à MV, a été apportée. En particulier, il a été montré qu'un modèle VAR (1) est équivalent à un M-V avec des marginales gaussiennes et des copules gaussiennes bivariées. La reprise de l'exemple théorique de Beare et Seo (2015) et une nouvelle application à un véritable jeu de données ont permis de mieux comprendre les développements et de rendre la procédure analytique plus concrète, respectivement.

Davantage d'applications et d'autres développements, y compris une méthode d'estimation plus compréhensible ou plus globale des paramètres d'une représentation M-V, sont nécessaires pour élargir la portée de la nouvelle approche. Étudier les processus chronologiques markoviens d'ordre supérieur, tels que le VAR( $p$ ) avec  $p \geq 2$ , et les relations entre l'approche classique et la représentation M-V demandera encore plus de réflexion et de travail, et cette remarque vaut pour d'autres modèles fréquemment utilisés tels le GARCH. Enfin, un test statistique futur, permettant de déterminer si une décomposition M-Vine représentant une série chronologique multivariée vaut le travail supplémentaire par rapport à une représentation du modèle VAR (1), serait fort utile.

# Bibliographie

- [ACFB09] Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insurance Math. Econom.*, 44(2) :182–198, 2009.
- [Aka74] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6) :716–723, 1974.
- [BC02] Tim Bedford and Roger M. Cooke. Vines—a new graphical model for dependent random variables. *Ann. Statist.*, 30(4) :1031–1068, 08 2002.
- [BD91] P.J. Brockwell and R.A. Davis. *Time Series : Theory and Methods*. Springer-Verlag, New York, 1991.
- [BJ76] G.E.P. Box and G.M. Jenkins. *Time series analysis : forecasting and control*. Holden-Day series in time series analysis and digital processing. Holden-Day, 1976.
- [Bol86] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3) :307–327, 1986.
- [BS15] Brendan K. Beare and Juwon Seo. Vine copula specifications for stationary multivariate Markov chains. *J. Time Series Anal.*, 36(2) :228–246, 2015.
- [Can17] Fabio Canova. *Vector Autoregressive Models : Specification, Estimation, Inference, and Forecasting*, chapter 2, pages 53–110. John Wiley and Sons, Ltd,



2017.

- [CF06] Xiaohong Chen and Yanqin Fan. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135(1-2) :125–154, 2006.
- [CH07] Song Xi Chen and Tzee-Ming Huang. Nonparametric estimation of copula functions for dependence modelling. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 35(2) :265–282, 2007.
- [Deh80] P. Deheuvels. Non parametric tests of independence. *Springer, Lecture Notes in Mathematics*, 821, 1980.
- [DL92] Pierre Dutilleul and Pierre Legendre. Lack of robustness in two tests of normality against autocorrelation in sample data. *Journal of Statistical Computation and Simulation*, 42(1-2) :79–91, 1992.
- [EMS02] Paul Embrechts, Alexander J. McNeil, and Daniel Straumann. *Correlation and Dependence in Risk Management : Properties and Pitfalls*, pages 176–223. Cambridge University Press, 2002.
- [End95] W. Enders. *Applied econometric time series*. John Wiley and Sons ltd., New York, 1995.
- [Eng82] Robert Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4) :987–1007, 1982.
- [FGS19] Mario Forni, Luca Gambetti, and Luca Sala. Structural vars and noninvertible macroeconomic models. *Journal of Applied Econometrics*, 34(2) :221–246, 2019.
- [FRW02] Jean-David Fermanian, Dragan Radulovic, and Marten Wegkamp. Weak convergence of empirical copula processes. (2002-06), 2002.

- [FV98] Edward W. Frees and Emiliano A. Valdez. Understanding relationships using copulas. *N. Am. Actuar. J.*, 2(1) :1–25, 1998.
- [Gat06] Jim Gatheral. *The Volatility Surface : A Practitioner’s Guide*. Wiley, 2006.
- [GGR95] C. Genest, K. Ghoudi, and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3) :543–552, 1995.
- [GOV12] Irene Gijbels, Marek Omelka, and Noel Veraverbeke. Multivariate and functional covariates and conditional copulas. *Electron. J. Statist.*, 6 :1273–1306, 2012.
- [GR93] Christian Genest and Louis-Paul Rivest. Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.*, 88(423) :1034–1043, 1993.
- [Ham94] J.D. Hamilton. *Time series analysis*. Princeton University Press, Princeton, New Jersey, 1994.
- [HH12] I. Hobaek Haff. Comparison of estimators for pair-copula constructions. *Journal of Multivariate Analysis*, 110 :91–105, 2012.
- [HH13] I. Hobaek Haff. Parameter estimation for pair-copula constructions. *Bernoulli*, 19 :462–491, 2013.
- [HS81] Harold V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and kronecker products : a review. *Linear and Multilinear Algebra*, 9(4) :271–288, 1981.
- [Joe96] Harry Joe. *Families of  $m$ -variate distributions with given margins and  $m(m-1)/2$  bivariate dependence parameters*, volume 28 of *Lecture Notes–Monograph Series*, pages 120–141. Institute of Mathematical Statistics, Hayward, CA, 1996.

- [Joe97] Harry Joe. *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1997.
- [Joe05] Harry Joe. Asymptotic efficiency of the two-stage estimation method for copula-based models. *J. Multivariate Anal.*, 94(2) :401–419, 2005.
- [JX96] H. Joe and J. Xu. The estimation method of inference functions for margins for multivariate models. *Technical Report*, 166, 1996.
- [KC06] Dorota Kurowicka and Roger M Cooke. *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, 2006.
- [KSS07] Gunky Kim, Mervyn J. Silvapulle, and Paramsothy Silvapulle. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics and Data Analysis*, 51(6) :2836–2850, 2007.
- [MK01] Dominique Drouet Mari and Samuel Kotz. *Correlation and Dependence*. EBL-Schweitzer. Imperial College Press, 2001.
- [MO13] Hugo Daniel Macedo and Jose Nuno Oliveira. Typing linear algebra : A biproduct-oriented approach. *Science of Computer Programming*, 78(11) :2160 – 2191, 2013. Special section on Mathematics of Program Construction (MPC 2010) and Special section on methodological development of interactive systems from Interaccion 2011.
- [Nel99] R.B. Nelsen. *An Introduction to Copulas*. Lecture notes in statistics. Springer, 1999.
- [Pat06] A. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47 :527–556, 2006.
- [Pri89] M.B. Priestley. *Spectral Analysis and Time Series*. Probability and Mathematical Statistics : a serie of monographs and textbooks. Academic Press, 1989.

- [RPS12] Bruno Remillard, Nicolas Papageorgiou, and Frederic Soustra. Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis*, 110 :30 – 42, 2012. Special Issue on Copula Modeling and Dependence.
- [Sch78] Gideon Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2) :461–464, 03 1978.
- [Sk159] Abe Sklar. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8 :229–231, 1959.
- [SS83] B Schweizer and A Sklar. *Probabilistic Metric Spaces*. Elsevier North-Holland, New York-Amsterdam-Oxford, 01 1983.
- [UY27] G. Udny Yule. On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer’s Sunspot Numbers. *Philosophical Transactions of the Royal Society of London Series A*, 226 :267–298, 1927.
- [Wal31] Sir Gilbert Walker. On periodicity in series of related terms. *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 131(818) :518–532, 1931.
- [Whi51] Peter Whittle. *Hypothesis testing in time series analysis*. Uppsala, Almqvist and Wiksells boktr., 1951.

# M-Vine decomposition and VAR(1) models

Étienne Begin<sup>a</sup>, Pierre Dutilleul<sup>b</sup>, Carole Beaulieu<sup>c</sup>, Taoufik Bouezmarni<sup>d,\*</sup>

<sup>a</sup>*Département de mathématiques, Université de Sherbrooke, Canada*

<sup>b</sup>*Department of Plant Science, McGill University, Canada*

<sup>c</sup>*Département de biologie, Université de Sherbrooke, Canada*

<sup>d</sup>*Département de mathématiques, Université de Sherbrooke, CIREQ, Canada*

---

## Abstract

The M-Vine (M-V) decomposition of 1st-order Vector Autoregressive [VAR(1)] models is detailed. We represent a VAR(1) with a Multivariate Gaussian Copula (MGC), build VAR(1) models using MGCs, decompose MGCs following M-V structures, and reconstruct an MGC from an M-V structure.

*Keywords:* M-Vine Decomposition; Vector Autoregressive Modeling; Multivariate Gaussian Copulas.

---

## 1. Introduction

Multivariate time series models have been applied in several domains, including finance, economics, and biology. Many parametric models have been proposed for analysing multiple time series, and the vector autoregressive (VAR) model appears to be the most used ([11]; [4]). Recently, copulas have become a popular method for modeling time series. Following [13], the multivariate distribution function,  $F$ , of a random vector  $\mathbf{X} = (X_1, \dots, X_k)^T$  (where  $^T$  denotes the transposition operator) can be decomposed into two components: the marginal distributions,  $F_1, \dots, F_k$ , and a copula function,  $C(\cdot)$ , which describes the dependence structure through the expression:

$$F(x_1, \dots, x_k) = C(F_1(x_1), \dots, F_k(x_k)).$$

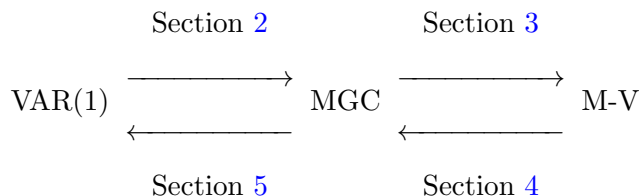
---

\*Corresponding author

The books of [10] and [8] provide ample details. More specifically, [5] proposed copula models for univariate time series. [12] extended the result of [5] to the case of multivariate time series. Selecting the copula function for high-dimensional data is all but straightforward. To ease this task, the vine decomposition was introduced for i.i.d. data by [7] and formally described by [3], [9], [1] and [6]. More recently, [2] proposed the M-Vine (M-V) structure for modeling multivariate time series, and showed that the VAR model of first order, or VAR(1), is equivalent to a model represented by a multivariate Gaussian copula (MGC) that can be decomposed with an M-V structure. In their article, [2] presented a theoretical bivariate example announced to be for a VAR(1) process and the M-V decomposition of its model. However, the M-V decomposition does not lead to the targeted VAR(1) model. That motivated us to study the relations existing between the VAR(1) process, MGC and the M-V decomposition or representation, and how to move from one representation to another.

Below, our article is organised as follows. In Section 2, the passage from VAR(1) to MGC is presented, with the required conditions on the VAR(1). In Section 3, we discuss a recursive calculation procedure developed to decompose an MGC into an M-V. Section 4 concerns the rebuilding of an MGC from its M-V. This rebuilding that we propose is based on another recursive procedure, mainly arising from the conditional theory for Gaussian distributions. Section 5 closes the loop by showing how to reshape an MGC in order to obtain its equivalent VAR(1). Lastly, in Section 6, the example in [2] is considered again.

Figure 1: Schematic presentation of sections



## 2. The multivariate Gaussian copula representation of a VAR(1) process

Let  $\mathbf{X}_t = (X_{1;t}, X_{2;t}, \dots, X_{k;t})^T$  be a  $k$ -variate VAR(1) time-series process defined by

$$\mathbf{X}_t = \boldsymbol{\mu} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t \quad \text{for } t \in \mathbb{N}, \quad (1)$$

where  $\boldsymbol{\mu}$  is a  $k$ -vector of constants,  $\Phi$  is the  $k \times k$  matrix of coefficients and  $\boldsymbol{\epsilon}_t, t \in \mathbb{N}$ , are i.i.d.  $k$ -variate normally distributed vectors with zero mean and variance-covariance matrix  $\boldsymbol{\Sigma}$ .

To find an MGC representation for this VAR(1), we calculate the joint distribution of  $(\mathbf{X}_t^T, \mathbf{X}_{t-1}^T)^T$ . We assume that  $(\mathbb{I} - \Phi)$  where  $\mathbb{I}$  is the identity matrix has an inverse and that  $\sum_{i=1}^t \Phi^{i-1} \boldsymbol{\Sigma} (\Phi^{i-1})^T$  converges to  $\boldsymbol{\Sigma}^*$  when  $t$  tends to infinity. Then, we can show that, that for large  $t$ ,  $(\mathbf{X}_t^T, \mathbf{X}_{t-1}^T)^T$  is a  $2k$ -variate normally distributed vector with mean  $\bar{\boldsymbol{\mu}}$  and variance-covariance matrix  $\bar{\boldsymbol{\Sigma}}$ :

$$\bar{\boldsymbol{\mu}} = \begin{pmatrix} (\mathbb{I} - \Phi)^{-1} \boldsymbol{\mu} \\ (\mathbb{I} - \Phi)^{-1} \boldsymbol{\mu} \end{pmatrix} \quad \text{and} \quad \bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Sigma}^* & \Phi \boldsymbol{\Sigma}^* \\ \boldsymbol{\Sigma}^* \Phi^T & \boldsymbol{\Sigma}^* \end{pmatrix}. \quad (2)$$

From the joint density of  $(\mathbf{X}_t^T, \mathbf{X}_{t-1}^T)^T$ , all the marginal densities are normal, i.e.  $X_{i;t} \sim N(\bar{\mu}_i, \bar{\Sigma}_{i,i})$  for  $i = 1, \dots, k$ . Moreover, the multivariate copula describing the dependence structure of  $(\mathbf{X}_t^T, \mathbf{X}_{t-1}^T)^T$  is a Gaussian copula. This MGC has  $k(k-1)/2$  parameters,  $\rho_{ih,jl}$ , which determine the correlations between the  $i$ th component of  $\mathbf{X}_{t+h-2}$  and the  $j$ th component of  $\mathbf{X}_{t+l-2}$  with  $i, j \in \{1, \dots, k\}$  and  $h, l \in \{1, 2\}$ ; note that for a VAR(1) model,  $\rho_{i1,j1}$  is equal to  $\rho_{i2,j2}$ . Using the MGC, the correlation parameters  $\rho_{ih,jl}$  can be derived from  $\bar{\boldsymbol{\Sigma}}$ :

$$\rho_{ih,jl} = \frac{\bar{\Sigma}_{(2-h)k+i, (2-l)k+j}}{\sqrt{\bar{\Sigma}_{i,i} \bar{\Sigma}_{j,j}}}. \quad (3)$$

## 3. The M-Vine decomposition of a multivariate Gaussian copula

[2] introduced the M-V decomposition of a multivariate first-order Markovian time series. The dependence structure is given by  $c_{11,12}$ , which is the density copula for the unconditional dependence between 11 and 12 where  $ih$  represents  $X_{i;t+h-2}$ , and  $c_{12,21|11}$  which is the conditional density copula of 12 and 21 given 11 (see the second tree). The construction of trees for an M-V leads to a decomposition

into bivariate copulas. [2] applied their methodology to VAR(1) models in a preliminary step, and stated without details that a VAR(1) model can be decomposed by M-Vine using Gaussian copulas.

We studied in detail and depth the M-V decomposition of VAR(1) models. Because we showed in Section 2 that a VAR(1) model is equivalent to an MGC in some conditions, we provide below the relation between M-V and MGC in order to decompose VAR(1) models by M-V.

In Section 2, we showed that a  $k$ -variate VAR(1) process,  $\mathbf{X}_t$ ,  $t \in \mathbb{N}$ , can be characterized by a  $2k$ -dimensional MGC, denoted  $c_\Omega$  where  $\Omega$  is a set of  $d = k(2k - 1)$  parameters. The M-V decomposes a  $2k$ -dimensional MGC into  $d$  bivariate Gaussian copulas with  $2k - 1$  unconditional copulas in the first tree. Hereafter, an unconditional copula is said to be direct in reference to direct correlation. All the other copulas, from the second to the  $(2k - 1)$ th tree, are conditional. Each bivariate copula corresponds to an edge of a tree in the M-V. Let  $E_w$  be the set of edges of the  $w$ th tree, so that  $e = c_{a_e, b_e | D_e} \in E_w$  where  $\{a_e, b_e\}$  is the conditioned set of  $e$  and  $D_e$  is the conditioning set of  $e$ . For example,  $E_2 = \{c_{11,31|21}, c_{21,12|11}, c_{11,22|12}, c_{12,32|22}\}$ .

Then, the decomposition of the copula  $c_\Omega$  proposed by [2] is given by:

$$c_\Omega = \prod_{w=1}^{2k-1} \prod_{e \in E_w} c_{a_e, b_e | D_e}. \quad (4)$$

By denoting  $f$  the joint density of  $(\mathbf{X}_{t-1}^T, \mathbf{X}_t^T)^T$  and  $f_{jl}$  the marginal of  $X_{j;t+l-2}$ , we have:

$$f = \left( \prod_{j \in \{1, \dots, k\}, l \in \{1, 2\}} f_{jl} \right) \left( \prod_{w=1}^{2k-1} \prod_{e \in E_w} c_{a_e, b_e | D_e} \right). \quad (5)$$

Considering only MGC models that are characterized by  $d$  parameters from the set  $\Omega$ , the parameters  $\rho_{ih,jl}$  represent the correlations between  $X_{i;t+h-2}$  and  $X_{j;t+l-2}$  where  $i, j \in \{1, \dots, k\}$  and  $h, l \in \{1, 2\}$ . With an MGC, the direct and conditional bivariate copulas of this M-V model are Gaussian. This implies that the parameters of the direct and conditional copulas of the M-V decomposition are the correlation and the conditional correlation, respectively. Direct correlations, i.e. the parameters of the first tree, are provided by the MGC. To obtain the parameters for the M-V, the conditional correlations associated with each edge of each of the trees, from the second to the  $(2k - 1)$ th, are required. To do that, the conditional theory for normal distributions is applied below.



We first define an operator for square matrices, denoted  $T'$ , which will simplify the developments. Let  $M = [m_{i,j}]_{i,j=1,\dots,k}$  be a  $k \times k$  square matrix. Then,  $M^{T'} = [m'_{i,j}]_{i,j=1,\dots,k}$ , where  $m'_{i,j} = m_{k+1-j,k+1-i}$  for  $i, j = 1, \dots, k$ . For example,

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{T'} = \begin{bmatrix} i & f & c \\ h & e & b \\ g & d & a \end{bmatrix}.$$

To optimise the recursive calculation used to obtain the parameters for the M-V, i.e. all the conditional correlations, we consider a matrix  $\Psi$  containing the parameters of the MGC, i.e. the direct correlations, organized as follows:

$$\Psi = \begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V}^T & \mathbf{U}^{T'} \end{pmatrix}, \quad (6)$$

where

$$\mathbf{U} = \begin{pmatrix} 1 & \rho_{(k-1)1,k1} & \cdots & \rho_{11,k1} \\ \rho_{(k-1)1,k1} & 1 & \cdots & \rho_{11,(k-1)1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{11,k1} & \rho_{11,(k-1)1} & \cdots & 1 \end{pmatrix} \text{ and } \mathbf{V} = \begin{pmatrix} \rho_{k1,12} & \cdots & \rho_{k1,k2} \\ \rho_{(k-1)1,12} & \cdots & \rho_{(k-1)1,k2} \\ \vdots & & \vdots \\ \rho_{11,12} & \cdots & \rho_{11,k2} \end{pmatrix}.$$

The theory of conditional Gaussian distributions tells us that with  $\mathbf{Z} = (\mathbf{X}^T, \mathbb{Y}^T)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , where  $\boldsymbol{\mu} = (\boldsymbol{\mu}_{\mathbf{X}}^T, \boldsymbol{\mu}_{\mathbb{Y}}^T)^T$  and  $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X},\mathbb{Y}} \\ \boldsymbol{\Sigma}_{\mathbb{Y},\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbb{Y}} \end{pmatrix}$ , the conditional mean and variance of  $\mathbf{X}$  given  $\mathbb{Y} = \mathbf{a}$  are defined by:

$$\boldsymbol{\mu}_{\mathbf{X}|\mathbb{Y}=\mathbf{a}} = \boldsymbol{\mu}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X},\mathbb{Y}} \boldsymbol{\Sigma}_{\mathbb{Y}}^{-1} (\mathbf{a} - \boldsymbol{\mu}_{\mathbb{Y}}) \quad \text{and} \quad \boldsymbol{\Sigma}_{\mathbf{X}|\mathbb{Y}=\mathbf{a}} = \boldsymbol{\Sigma}_{\mathbf{X}} - \boldsymbol{\Sigma}_{\mathbf{X},\mathbb{Y}} \boldsymbol{\Sigma}_{\mathbb{Y}}^{-1} \boldsymbol{\Sigma}_{\mathbb{Y},\mathbf{X}}. \quad (7)$$

Since correlations are invariant to scale and origin change of marginals, we assume for the remainder of the paper that the marginal distributions are standard Gaussian distributions in the calculation of conditional correlations (and that of direct correlations in Section 4).

To calculate the entries of the conditional correlation matrix, we use (7) in a recursive way:

$$\psi^{j+1,b} = \text{diag} \left( \psi_1^{j,b} - \psi_2^{j,b} \left( \psi_2^{j,b} \right)^T \right)^{-1/2} \left( \psi_1^{j,b} - \psi_2^{j,b} \left( \psi_2^{j,b} \right)^T \right) \text{diag} \left( \psi_1^{j,b} - \psi_2^{j,b} \left( \psi_2^{j,b} \right)^T \right)^{-1/2} \quad (8)$$

where, for  $b = 2, \dots, 2k - 1$  and  $j = 0, \dots, 2k - b$ ,  $\psi_1^{j,b}$  is equal to matrix  $\psi^{j,b}$  without its  $b$ th row and  $b$ th column, and  $\psi_2^{j,b}$  is the  $b$ th column of  $\psi^{j,b}$  without its  $b$ th term (i.e. the entry equal to 1 in the column), with  $\psi^{0,b} = \Psi$ ; and  $\text{diag}(\cdot)$  is the diagonal operator which returns a matrix with only diagonal non-null elements, equal to those of the square matrix used as argument.

Equation (8) is very practical, and provides all the required conditional correlations for an M-V decomposition in the case of a  $k$ -variate first-order Markovian time series. All the parameters of the M-V can be obtained from an MGC in the following two steps:

Step 1: The initial matrix of correlation parameters,  $\Psi$ , provides the direct correlations of the M-V decomposition.

Step 2: For  $b$  ranging from 2 to  $2k - 1$ ,

1.  $\psi^{0,b}$  is set at  $\Psi$ .
2. For  $j$  ranging from 1 to  $2k - b$ , (a) Calculate  $\psi^{j,b}$  using (8); (b)  $\psi_{(b-1),b}^{j,b}$  is a parameter of the M-V.

Eventually, all the M-V parameters are obtained sequentially:  $\rho_{(k-1)1,k1}, \rho_{(k-2)1,(k-1)1}, \dots, \rho_{21,11}, \rho_{11,12}, \rho_{12,22}, \dots, \rho_{(k-1)2,k2}, \rho_{(k-2)1,k1|(k-1)1}, \dots, \rho_{11,k1|(k-1)1,(k-2)1,\dots,21}, \rho_{k1,12|(k-1)1,\dots,11}, \dots, \rho_{k1,k2|(k-1)1,\dots,11,12,\dots,(k-1)2}, \rho_{(k-3)1,(k-1)1|(k-2)1}, \dots, \rho_{(k-2)2,k2|(k-1)2}$ . Hence, all the parameters of the bivariate Gaussian copulas involved in the M-V decomposition can be deduced from the MGC model. Therefore, a VAR(1) model can be represented with an M-V since we showed that a VAR(1) model can be rewritten as an MGC in Section 2.

#### 4. Reconstruction of a multivariate Gaussian copula from its M-Vine decomposition

In this section, we show the reverse, i.e. how to obtain an MGC from its M-V decomposition, the ultimate goal being to build a VAR(1) model from an M-V decomposition. Therefore, we consider an

M-V with Gaussian marginals and Gaussian bivariate copulas because if only one copula or marginal distribution from an M-V is not Gaussian, it will be impossible to reconstruct an MGC and the underlying model for the multivariate time series will not be VAR(1).

Consider an M-V for a  $k$ -variate first-order Markovian time series with all marginals and bivariate copulas of the decomposition being Gaussian, and assume that all the parameters are known; we shall see in Section 6 how the parameters can be estimated. The M-V model is characterized by  $d = k(2k - 1)$  bivariate Gaussian copulas. The parameters of these Gaussian copulas represent the correlation between two time-series processes:  $2k - 1$  of those parameters correspond to the direct correlations from the first tree of the decomposition, and the rest corresponds to conditional correlation.

For an MGC model, the parameters are the  $d$  direct correlations between the series  $X_{i,t+h-2}$  and  $X_{j,t+l-2}$  with  $ih \neq jl$ ,  $i, j \in \{1, \dots, k\}$  and  $h, l \in \{1, 2\}$ . Therefore, we calculate the  $(k - 1)(2k - 1)$  unknown direct correlations via the parameters of the M-V. Consider  $\Psi$ , the matrix containing all the MGC parameters; see (6) in Section 3. The only known elements of  $\Psi$  are the  $(2k - 1)$  direct correlations given by the M-V and positioned just above the diagonal. Loosely speaking, we call ‘second diagonal’ the set of these elements, those just above them ‘third diagonal’, and so on. For example, the ‘ $(2k)$ th diagonal’ of  $\Psi$  is given by  $\rho_{k1,k2}$ ; the ‘first diagonal’ (the classical one) is made of ones.

To calculate  $\rho_{ih,jl}$  (the correlation between  $X_{i,t+h-2}$  and  $X_{j,t+l-2}$ ) on the  $q$ th diagonal of  $\Psi$  for  $q \geq 3$  (without loss of generality,  $i > j$ ), consider the symmetric square sub-matrix of  $\Psi$ :

$$\Psi^{(ih,jl)} == \left[ \begin{array}{c|c|c} 1 & \Psi_{21}^{(1)} & \rho_{ih,jl} \\ \hline \Psi_{12}^{(1)} & \Psi_{22} & \Psi_{12}^{(2)} \\ \hline \rho_{ih,jl} & \Psi_{21}^{(2)} & 1 \end{array} \right].$$

Define:

$$\begin{aligned}\Psi_1^{(ih,jl)} &= \begin{pmatrix} 1 & \rho_{ih,jl} \\ \rho_{ih,jl} & 1 \end{pmatrix}, & \Psi_2^{(ih,jl)} &= \Psi_{22}, \\ \Psi_3^{(ih,jl)} &= \begin{pmatrix} \Psi_{12}^{(1)} & \Psi_{12}^{(2)} \end{pmatrix}, & \Psi_4^{(ih,jl)} &= \begin{pmatrix} \Psi_{21}^{(1)} \\ \Psi_{21}^{(2)} \end{pmatrix}.\end{aligned}$$

Note that the partitioning of  $\Psi^{(ih,jl)}$  above is for  $h, l = 1$ . As for  $h, l = 2$ , the relevant sub-matrix is given by  $(\Psi^{(ih,jl)})^{T'}$ . Sub-matrix  $\Psi^{(ih,jl)}$  when  $h \neq l$  can be derived similarly.

As in Section 3, we assume hereafter that the marginals are standard Gaussian distributions. Then,  $\Psi^{(ih,jl)}$  represents the matrix of variances and covariances for  $(X_{i;t+h-2}, X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}, X_{j;t+l-2})^T$ , which follows a multivariate standard Gaussian distribution. Moreover, to calculate the parameters of the  $q$ th diagonal, assume that all the parameters in  $\Psi$  up to the  $(q-1)$ th diagonal are known. From (7), it follows:

$$\begin{bmatrix} \sigma_{ih|D}^2 & \rho_{ih,jl|D} \sigma_{ih|D} \sigma_{jl|D} \\ \rho_{ih,jl|D} \sigma_{ih|D} \sigma_{jl|D} & \sigma_{jl|D}^2 \end{bmatrix} = \begin{bmatrix} 1 - \beta_{11}^{(ih,jl)} & \rho_{ih,jl} - \beta_{12}^{(ih,jl)} \\ \rho_{ih,jl} - \beta_{12}^{(ih,jl)} & 1 - \beta_{22}^{(ih,jl)} \end{bmatrix}, \quad (9)$$

where  $D$  is the set  $\{(i-1)h, \dots, (j+1)l\}$ ,  $\sigma_{ih|D}^2$  and  $\sigma_{jl|D}^2$  are the conditional variances of the random variables  $X_{i;t+h-2}|X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}$  and  $X_{j;t+l-2}|X_{i-1;t+h-2}, \dots, X_{j+1;t+l-2}$ , respectively;  $\rho_{ih,jl|D}$  is one of the  $2k-q+1$  parameters of the  $q$ th tree of the M-V and  $\beta^{(ih,jl)} = \Psi_{12}^{(ih,jl)} (\Psi_{22}^{(ih,jl)})^{-1} \Psi_{21}^{(ih,jl)}$

is the symmetric  $2 \times 2$  matrix  $\begin{pmatrix} \beta_{11}^{(ih,jl)} & \beta_{12}^{(ih,jl)} \\ \beta_{12}^{(ih,jl)} & \beta_{22}^{(ih,jl)} \end{pmatrix}$ .

Equation (9) leads to a system of three equations for three unknown parameters:

$$\sigma_{ih|D}^2 = 1 - \beta_{11}^{(ih,jl)}, \sigma_{jl|D}^2 = 1 - \beta_{22}^{(ih,jl)} \text{ and } \rho_{ih,jl|D} \sigma_{ih|D} \sigma_{jl|D} = \rho_{ih,jl} - \beta_{12}^{(ih,jl)}. \quad (10)$$

From (10), we obtain:

$$\rho_{ih,jl} = \rho_{ih,jl|D} \sqrt{(1 - \beta_{11}^{(ih,jl)}) (1 - \beta_{22}^{(ih,jl)})} + \beta_{12}^{(ih,jl)}. \quad (11)$$

Accordingly, all the direct correlations of  $\Psi$  can be obtained from the M-V. The recursive procedure below summarizes the developments that provide all the parameters for an MGC from an M-V with Gaussian marginals and bivariate copulas:

1. Let  $\Psi$  be the symmetric  $2k \times 2k$  matrix of parameters of the MGC as in (6) and  $\Omega'$ , the set of parameters of the M-V.
2. The terms on the second diagonal of  $\Psi$  are obtained from the parameters of the first tree of the M-V in  $\Omega'$ .
3. For  $q$  ranging from 3 to  $2k$ ,
  - (a) Consider the  $q$ th diagonal of  $\Psi$ , composed of  $2k - q + 1$  parameters.
  - (b) For  $u$  ranging from 1 to  $2k - q + 1$ ,
    - i. Consider the  $u$ th element on the  $q$ th diagonal of  $\Psi$  that corresponds to an unknown parameter  $\rho_{ih,jl}$  with:

$$\begin{cases} i = k - u + 1 & \text{and} & h = 1 & \text{if } u \leq k \\ i = u - k & \text{and} & h = 2 & \text{otherwise} \end{cases}$$

and

$$\begin{cases} j = k - u - q + 2 & \text{and} & l = 1 & \text{if } q + u \leq k + 1 \\ j = q + u - k - 1 & \text{and} & l = 2 & \text{otherwise.} \end{cases}$$

- ii. Calculate  $\beta^{(ih,jl)}$  as described above.
  - iii. Using (11) and  $\rho_{ih,jl|i-1h,\dots,j+1l} \in \Omega'$ , calculate  $\rho_{ih,jl}$ .
  - iv. Enter the value found for  $\rho_{ih,jl}$  in the relevant entry of  $\Psi$ .
4. All the entries of matrix  $\Psi$  are calculated.

## 5. The VAR(1) behind a multivariate Gaussian copula representation

In the previous sections, we studied the MGC representation of a VAR(1), the M-V decomposition of an MGC model and the building of an MGC from its M-V. In this section, we find the underlying VAR(1) process for a given MGC. Let  $\mathbf{X}_t = (X_{1;t}, X_{2;t}, \dots, X_{k;t})^T$  be a  $k$ -variate first-order Markovian process represented by an MGC, i.e. the  $k$  Gaussian marginals and all the parameters, which are direct

correlations between series at the same time ( $t$ ) and at crossed times ( $t$  and  $t - 1$ ), are thus identified. The mean and the variance of  $X_{i;t+h-2}$  are respectively denoted  $\mu_i$  and  $\sigma_i^2$ , with  $i \in \{1, \dots, k\}$  and  $h \in \{1, 2\}$ . Then, the joint distribution of  $(\mathbf{X}_{t-1}^T, \mathbf{X}_t^T)^T$  is given by:

$$\begin{pmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_t \end{pmatrix} \sim N \left( \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2^T & \boldsymbol{\Sigma}_1 \end{pmatrix} \right), \quad (12)$$

where  $\boldsymbol{\mu}$  is the vector mean of  $\mathbf{X}$  and  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  are the variance-covariance matrix.

To obtain the VAR(1) model, we need the distribution of  $\mathbf{X}_t | \mathbf{X}_{t-1}$ . From (12) and the conditional theory for Gaussian distributions, we have:

$$\boldsymbol{\mu}_{\mathbf{X}_t | \mathbf{X}_{t-1}} = \bar{\boldsymbol{\mu}} + \Phi \mathbf{X}_{t-1} \quad (13)$$

and

$$\boldsymbol{\Sigma}_{\mathbf{X}_t | \mathbf{X}_{t-1}} = \bar{\boldsymbol{\Sigma}}. \quad (14)$$

Hence, from (13) and (14), we obtain  $\mathbf{X}_t | \mathbf{X}_{t-1} \sim N(\bar{\boldsymbol{\mu}} + \Phi \mathbf{X}_{t-1}, \bar{\boldsymbol{\Sigma}})$ , and the VAR(1) representation of  $\mathbf{X}_t$  is given by:  $\mathbf{X}_t = \bar{\boldsymbol{\mu}} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$ , where  $\boldsymbol{\epsilon}_t \sim N(0, \bar{\boldsymbol{\Sigma}})$ ,  $\bar{\boldsymbol{\mu}} = (\mathbb{I} - \boldsymbol{\Sigma}_2^T \boldsymbol{\Sigma}_1^{-1}) \boldsymbol{\mu}$ ,  $\Phi = \boldsymbol{\Sigma}_2^T \boldsymbol{\Sigma}_1^{-1}$  and  $\bar{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_2$ .

It follows that a time-series process represented with an MGC can be rewritten as a VAR(1). Also, it must be noted that there is no required condition on the MGC to make the change of representation, which implies that each MGC has an equivalent VAR(1).

## 6. Theoretical example revisited

[2] discussed the relation between VAR(1) and M-V. With minimal explications, they suggested that a bivariate first-order Markovian time-series process represented as a VAR(1) model,  $\mathbf{X}_t = \boldsymbol{\mu} + \Phi \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$ , with  $\boldsymbol{\epsilon}_t \sim \text{i.i.d. } N(0, \boldsymbol{\Sigma})$  and the parameters:

$$\boldsymbol{\mu} = \begin{bmatrix} 3.5 \\ -0.5 \end{bmatrix}, \quad \Phi = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \quad (15)$$

is equivalent to an M-V modeling in which the marginals are Gaussian  $N(5, 1.6638^2)$  for  $X_{1t}$  and  $N(0, 1.2377^2)$  for  $X_{2t}$ , and the parameters are:

$$\begin{aligned}\rho_{11,21} &= 0.7243, & \rho_{11,12} &= 0.4078, & \rho_{12,22} &= 0.7243, \\ \rho_{21,12|11} &= 0.1123, & \rho_{11,22|12} &= -0.1406, & \rho_{21,22|11,12} &= -0.0144.\end{aligned}\tag{16}$$

Below, we consider the VAR(1) model (15) and change its representation from VAR(1) to MGC and then from MGC to M-V. Reversely, we begin with the M-V representation (16) of [2], to obtain a VAR(1) model via the MGC model.

Some regular conditions must be satisfied to write a VAR(1) model as an MGC (Section 2). All of them are met for the VAR(1) model (15). With (2) and (3) from Section 2, the MGC model is obtained with marginals  $X_{1t} \sim N(5, 1.1053^2)$  and  $X_{2t} \sim N(0, 1.017^2)$ , and the following parameters:

$$\begin{aligned}\rho_{11,21} &= 0.5219, & \rho_{11,12} &= 0.3960, & \rho_{12,22} &= 0.5219, \\ \rho_{21,12} &= 0.3406, & \rho_{11,22} &= 0.1609, & \rho_{21,22} &= 0.1567.\end{aligned}\tag{17}$$

From the MGC model (17) and the two steps defined in Section 3, we find an M-V decomposition with marginals  $X_{1t} \sim N(5, 1.1053^2)$  and  $X_{2t} \sim N(0, 1.017^2)$ , and the parameters:

$$\begin{aligned}\rho_{11,21} &= 0.5219, & \rho_{11,12} &= 0.3960, & \rho_{12,22} &= 0.5219, \\ \rho_{21,12|11} &= 0.1709, & \rho_{11,22|12} &= -0.0585, & \rho_{21,22|11,12} &\approx 0,\end{aligned}\tag{18}$$

which is not the M-V model (16) proposed by [2].

Furthermore, consider the suggested M-V model (16). From this M-V representation, we obtain using the recursive procedure from Section 4, an MGC model with marginals  $N(5, 1.6638^2)$  for  $X_{1t}$  and  $N(0, 1.2377^2)$  for  $X_{2t}$ , and the following parameters:

$$\begin{aligned}\rho_{11,21} &= 0.7243, & \rho_{11,12} &= 0.4078, & \rho_{12,22} &= 0.7243, \\ \rho_{21,12} &= 0.3661, & \rho_{11,22} &= 0.2068, & \rho_{21,22} &= 0.1974.\end{aligned}\tag{19}$$

Following the procedure from Section 5, we obtain a VAR(1) model from the MGC model (19) with parameters:

$$\boldsymbol{\mu} = \begin{bmatrix} 3.5 \\ -0.5 \end{bmatrix}, \quad \boldsymbol{\Phi} = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 2.2787 & 1.3033 \\ 1.3033 & 1.4591 \end{bmatrix}, \quad (20)$$

which is not the VAR(1) model (15) proposed by [2].

It follows that [2] obtained (16), whereas the true equivalent M-V model is (18) above. It can also be observed that the parameter values that we find for the VAR(1) model equivalent to the M-Vine decomposition proposed by [2] [see (20) above] are different from those of the VAR(1) model reported in their article [(15) above].

## Acknowledgements

T. Bouezmarni gratefully acknowledges the research support of the Natural Sciences and Engineering Research Council of Canada (NSERC) and Fonds de recherche du Québec - Nature et technologies (FRQNT). Étienne Begin gratefully acknowledges a scholarship from NSERC CREATE Training Program to meet the technological, environmental and socio-economic challenges of crop production.

## References

- [1] Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insurance Math. Econom.*, 44(2):182–198, 2009.
- [2] Brendan K. Beare and Juwon Seo. Vine copula specifications for stationary multivariate Markov chains. *J. Time Series Anal.*, 36(2):228–246, 2015.
- [3] Tim Bedford and Roger M. Cooke. Vines—a new graphical model for dependent random variables. *Ann. Statist.*, 30(4):1031–1068, 2002.
- [4] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, 1990.
- [5] Xiaohong Chen and Yanqin Fan. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *J. Econom.*, 135(1-2):125–154, 2006.



- [6] Claudia Czado. Pair-copula constructions of multivariate copulas. In *Workshop on Copula Theory and its Applications*, volume 198, pages 93–109. Springer, Dordrecht, 2010.
- [7] Harry Joe. *Families of  $m$ -variate distributions with given margins and  $m(m-1)/2$  bivariate dependence parameters*, volume 28 of *Lecture Notes–Monograph Series*, pages 120–141. Institute of Mathematical Statistics, Hayward, CA, 1996.
- [8] Harry Joe. *Dependence modeling with copulas*. Taylor and Francis, New York, 2016.
- [9] Dorota Kurowicka and Roger M Cooke. *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, Chichester, 2006.
- [10] Roger B. Nelsen. *An Introduction to Copulas*. Lecture notes in statistics. Springer, Berlin, 1999.
- [11] Maurice B. Priestley. *Spectral analysis and time series*. Academic Press, London, 1981.
- [12] Bruno Remillard, Nicolas Papageorgiou, and Frederic Soustra. Copula-based semiparametric models for multivariate time series. *J. Mult. Anal.*, 110:30–42, 2012. Special Issue on Copula Modeling and Dependence.
- [13] Abe Sklar. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959.